

https://dspace.org/

RB1 and Beyond: Determining genetic causes of retinoblastoma in South African patients

Beukman, Danielle

2024

University of the Witwatersrand, Johannesburg

https://hdl.handle.net/10539/42965

Beukman, Danielle. (2024). RB1 and Beyond: Determining genetic causes of retinoblastoma in South African patients [Master's dissertation PhD thesis, University of the Witwatersrand, Johannesburg]. WireDSpace.

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

RB1 AND BEYOND:

DETERMINING GENETIC CAUSES OF RETINOBLASTOMA IN SOUTH AFRICAN PATIENTS

Danielle Beukman
Student number: 719425

Supervisors:
Dr. Lindie Lamola
Ms. Nkateko Mayevu

A research report in the format of a “submissible” paper submitted to the Faculty of Health Sciences,
University of the Witwatersrand, Johannesburg.
This research report is required for partial fulfilment of the requirements for the degree of Master of Science in Medicine Genomic Medicine



Johannesburg, South Africa
25 February 2024

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

30 November 2023

Re: Research report submitted to the Faculty of Health Sciences by Ms. Danielle Beukman in partial fulfilment of the requirements for the degree of Master of Science in Medicine (Genomic Medicine)
I would like to thank you for agreeing to examine my research report titled "RB1 and beyond: Determining genetic causes of retinoblastoma in South African patients"

The candidate, Ms. Danielle Beukman, is submitting her research report in the format of a "submissible" format. The research report contributes 33% towards the final mark of the degree. The candidate and her supervisors have chosen to submit the paper to the journal "Human Genetics" published by Springer. The submissible paper attached is therefore written in accordance with the author guidelines of the above-mentioned journal. These guidelines have been attached as a website link for your reference. The only deviation from these guidelines is that the manuscript has been presented as one complete document (including tables, figures, and list of abbreviations) to make marking easier, rather than being submitted separately.
The research protocol is attached as an appendix for the purposes of providing an extended literature review and further context to the study. The protocol has already been assessed and approved by the Faculty of Health Sciences 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,



Student: Ms. Danielle Beukman



Supervisor: Dr Lindie Lamola



Co-supervisor: Ms. Nkateko Mayevu

Declaration

I, Danielle Beukman, declare that this research report, in the form of a “submissible” paper, is my own, unaided work. It has not been submitted before for any degree or examination at any other University.

This research report is required for the Degree of Master of Science Genomic Medicine at the University of the Witwatersrand, Johannesburg.

Danielle Beukman
WITS student number 719425



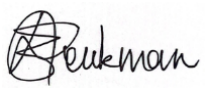
Signature of candidate, Danielle Beukman

Signed on 26 February 2024, in Johannesburg, RSA

Contribution of the candidate to the paper

Declaration: Student's contribution to article(s) and agreement of co-authors

I, Danielle Beukman, student number 719425, declare that this Research Report, entitled "RB1 and Beyond: Determining genetic causes of Retinoblastoma in South African patients" is my own work and that I made significant contributions towards the research findings presented in this paper, intended to be submitted for publication below.



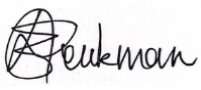
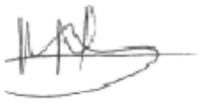
Signature of Student Date: 15 November 2023

Signature of Primary Supervisor Dr. L.Lamola

Date: 16 November 2023

Agreement by co-authors

By signing this declaration, the co-authors listed below agree to the use of the research d fini ngs in tended for a published article by a student as part of their Research Report.

Authors	Name	Signature	Date
1st	Ms. Danielle Beukman		15 November 2023
2nd	Dr. Lindiwe Lamola		16 November 2023
3rd	Ms. Nkateko Mayevu		16 November 2023

Dedication

I dedicate this research report to my family, in particular my beloved mother, Margaret and elder sister, Natasha. Your everlasting support has immeasurable value.

Abstract

Retinoblastoma is the most common solid tumour of the retina, affecting mainly paediatric patients. Although loss-of-function germline variants the *RB1* gene is the tumour suppressor gene most often associated with heritable retinoblastoma, two contradictory research findings complicate the assumption that a pathogenic germline variant in *RB1* will lead to the development of heritable retinoblastoma. Firstly, heritable retinoblastoma has been observed in cohorts with biallelic wild-type *RB1* genes, secondly, biallelic loss-of-function is not always sufficient to lead to retinoblastoma tumorigenesis. By investigating germline variants in additional cancer predisposition genes to discover candidate driver genes in heritable retinoblastoma, the full germline mutational spectrum of heritable retinoblastoma can be established. In this study we investigate sequence germline variants in cancer predisposition genes beyond *RB1* in South African retinoblastoma patients with an African ancestry by performing whole exome sequencing on eight patients diagnosed with retinoblastoma. Whole exome sequencing data for genes previously associated with retinoblastoma was subjected to variant annotation by means of a variant effect predictor and filtering criteria applied manually. Variant classification was performed according to guidelines proposed by The American College of Medical Genetics and Genomics and the Association of Molecular Pathology. The highest degree of variant classification was variants of unknown significance. The absence of sequence variants capable of describing the retinoblastoma phenotype in this cohort demonstrates the necessity of looking beyond *RB1*, combining next generation sequencing with copy number analysis pipelines and additional genome mapping technologies capable of detecting structural variants. The detection of sequence variants previously reported to be possibly damaging but now considered polymorphisms highlights the importance of revisiting variant classification. In closing, this study adds genomic information from patients with a South African ancestry to mounting genomic data, ensuring that previously underrepresented populations can also benefit from future cancer predisposition and precision medicine research.

Acknowledgements

I would like to give special thanks to my supervisors, Dr. Lindie Lamola and Ms. Nkateko Mayevu. Your guidance, support and advice made me grow ever curious, but never defeated. I appreciate your patience and hard work. Thank you for allowing me to engage in the PeCan Project, an inspiring research venture that is now a cause close to my heart.

Thank you to the WITS Division of Human Genetics academics. Professor Krause, Ms. Monica Araujo, Doctor L. Lamola, Doctor R. Kerr and Doctor Z. Lombard, and Doctor F. Baine-Savanhu, your passion for research and teaching is remarkable. Thank you all for the time you dedicated to us. Your enthusiasm for advancing genetic and genomic research in South Africa is infectious.

I feel honored to have shared this year with my classmates, Anréé, Ingrid and Racilya, your friendship made every Wednesday worth it! I cherish every moment. Dylan, Thandeka, Phophi and Gabriella, thank you for being there when I needed advice.

Most importantly, to God and my family, thank you for keeping the rent low and spirits high.

To my mom Kleintjie, thank you for your unconditional support and the endless supply of Earl Grey tea. Dankie ma.

Table of Contents

Declaration.....i
Contribution of the candidate to the paper ii
Dedication iii
Abstract iv
Acknowledgements v
Table of Contents.....vi
List of Appendices vii
List of Figures.....viii
List of Tables.....ix
Table of Abbreviations x
Research report in the form of a "submissible" paper.....xi
Title page.....xi
1. **Abstract for journal submission** 1
2. **Background**.....2
2.1. Retinoblastoma in South Africa.....2
2.2. Future implications of a definitive genetic diagnosis.....2
2.3. The genetic landscape of RB.....3
2.4. Locus heterogeneity observed in RB predisposition.....4
3. **Materials and Methods**.....7
3.1. Ethical considerations and Consent 7
3.2. Clinical Presentation and patient demographics.....7
3.3. Whole exome sequencing 8
3.3.1. DNA sample preparation and sample quality control.....8
3.3.2. Library generation 8
3.3.3. Templating and sequencing.....8
3.4. Variant interpretation.....9
4. **Results** 10
4.1. Clinical presentation of patients.....10
4.2. Variant annotation.....10
4.3. *RB1* sequence variants 11
4.4. Variants of uncertain significance 11
4.5. Shared variants.....14
5. **Discussion** 15
5.1. *RB1* variants 15
5.2. Shared variants 16
5.3. Variants of uncertain significance 17
5.4. Future prospects and limitations.....18
6. **Conclusion** 19
7. **References**.....20
Appendices.....29

List of Appendices

Appendix A:

Human Research Ethics Committee (Medical), University of the Witwatersrand Ethics Certificate (M180855)

Appendix B:

Human Research Ethics Committee (Medical), University of the Witwatersrand Ethics Certificate (M230782)

Appendix C:

RB1 sequence variants in all patients

Appendix D

Extended list of variants of interest

Appendix E

BAM images to evaluate variant coverage

Appendix F:

Information, Consent and Assent forms

Appendix G:

Title Approval and Approved Research Protocol

Appendix H:

Plagiarism declaration and Turnitin report

Appendix I:

Journal Author Guidelines for Submissible format

List of Figures

Figure 1:
Distribution of annotated variants……10

Figure 2:
Distribution of variants in coding regions ……11

List of Tables

Table 1:
Genes implicated in the aetiology of RB development ……6

Table 2:

Clinical presentation of RB patients……7

Table 3:
Subset of variants of interest as observed per patient ……13

Table 4:
Shared variants ……14

Table of Abbreviations

Abbreviation	Meaning
ACMG/AMP	American College of Medical Genetics and Genomics and Association for Molecular Pathology
BAM	Binary alignment map
CADD	Combined Annotation-Dependent Depletion
CMA	Chromosomal microarray
CNV	Copy number variation
HREC Medical	Human Research Ethics Committee (Medical)
IGV	Integrative genomics viewer
Indels	Insertions and Deletions
LOVD	Leiden Open Variation Database
MAF	Minor Allele Frequency
miRNA	Micro RNA
MLPA	Multiplex ligation-dependent probe amplification
PeCan	Paediatric Cancer
POLYPHEN2	Polymorphism Phenotyping version 2
RB	Retinoblastoma
RB1	Retinoblastoma 1 tumor suppressor gene
SNV	Single Nucleotide Variants
SIFT	Sort Intolerable from Tolerable
SANCR	South African National Cancer Registry
TSG	*Tumor suppressor gene*
VCF	Variant Call files
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 the Journal Human Mutation

RB1 and Beyond: Determining genetic causes of Retinoblastoma in South African patients

Danielle Beukman[1], Nkateko Mayevu[1], Lindie Lamola[1]

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

Corresponding Author:

Danielle Beukman

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

[T]: 084 583 4985 [E]: beukmandanielle@gmail.com / 719425@students.wits.ac.za

1. Abstract for Journal Submission

Background: Retinoblastoma is the most common solid tumour of the retina, affecting mainly paediatric patients. Although loss-of-function germline variants the *RB1* gene is the tumour suppressor gene most often associated with heritable retinoblastoma, two confounding research findings contradict the direct genotype phenotype relationship. First, heritable retinoblastoma has been observed in cohorts with biallelic wild-type *RB1* genes, secondly, biallelic loss-of-function is not always sufficient to lead to retinoblastoma tumorigenesis. By investigating germline variants in additional cancer predisposition genes beyond *RB1* allows for discovery of candidate driver genes in heritable retinoblastoma. In this study we investigate sequence germline variants in cancer predisposition genes beyond *RB1* in retinoblastoma patients with a South African ancestry.

Methods: Whole exome sequencing with single nucleotide resolution was performed to establish germline sequence variants. Variant annotation, prioritization and filtering was executed to detect germline variants in cancer predisposition genes previously proposed as candidate genes in retinoblastoma malignancy and paediatric cancer predisposition syndromes. Variant prioritization excluded alleles with a minor allele frequency above 5% and variants predicted to have no impact on a protein level due to no change in the amino acid sequence and protein structure. Variants shared between individuals and variants of interest were classified according to ACMG/AMP guidelines.

Results: A total of 353 655 variant were annotated, 4538 of which were novel. After collating functional data, previously reported and reviewed *RB1* variants, all sequence variants in *RB1* are likely benign, (NM_000321.3:c.1574C>G, NP_000312.2:p.Ala525Gly and NM_000321.3:c.45_53del NP_000312.2:p.Ala16_Ala18del) amending outdated variants incorrectly classified as likely pathogenic. The highest degree of variant classification in genes previously associated with retinoblastoma are variants of unknown significance.

Conclusion: The absence of sequence variants capable of describing the retinoblastoma phenotype in this cohort demonstrates the necessity of combining next generation sequencing with analysis pipelines and additional genome mapping to detect structural variants. This study adds genomic information from patients with a South African ancestry, an underrepresented population.

Keywords: germline variation, cancer predisposition genes, whole exome sequencing, *RB1* gene, retinoblastoma, sequence variants

2. Background

2.1. Retinoblastoma in South Africa

Retinoblastoma (RB) is the most commonly occurring solid intraocular cancer, developing in childhood. The South African National Cancer Registry (SANCR) documented an average RB incidence of 1 in 21 641 live births for the time elapsed between 2004 and 2018 (https://www.nicd.ac.za/centres/national-cancer-registry). South African patients have a later mean age of diagnosis (2.5 years), compared to an earlier mean age of diagnosis of 1.2 years in high income countries (Stuart et al., 2022). Routine cancer surveillance is poor in South Africa due to an overburdened public health care system. As an early age of diagnosis is the strongest indicator of an improved prognosis, a later mean age of diagnosis puts South African children suffering from paediatric cancers in a disadvantageous position (Stuart et al., 2022).

Recent efforts have been made to address South Africa's later mean age of cancer diagnosis compared to global averages. The overall survival rate for RB patients in South Africa increased from 52% in 1994 (Hesseling et al., 1995) to 86.7% twenty years later (Ndlovu et al., 2023). Limiting disease for RB exhibits a 100% overall survival rate, emphasizing the importance of early detection. Children with a black South African ancestry have a disproportionately high incidence of RB compared to other racial groups (Ndlovu et al., 2023).

The majority (60%) of RB cases are sporadic and unilateral (tumors affecting a single eye), occurring at a mean age of 24 months. The remaining 40% of RB cases are heritable, bilateral (tumors affecting both eyes) RB, with a significantly earlier mean age of onset of 12 months (Bouchoucha et al., 2023; Knudson, 1971). A small proportion of RB patients present with pineoblastoma in conjunction with bilateral RB, referred to as trilateral RB. The presence of a constitutional (germline) mutation in *RB1* can either be inherited from a parent or occur de novo at an early stage of embryonic development, a mutation occurring at a later stage of embryonic development will result in mosaicism (Lin & Chintagumpala, 2021; Rushlow et al, 2009).

2.2. Future implications of a definitive genetic diagnosis

The detection of a disease-causing germline variant and a mutation signature can aid in treatment planning and risk assessment. Patients with a pathogenic germline variant in *RB1* have a predilection to developing further malignancies (Ketteler et al., 2020; Shinohara et al., 2014) and a recurrence risk

in offspring. A literature search into genetic variants predisposing black South African individuals to RB malignancy proved futile, revealing a gap in our understanding of hereditary and paediatric cancers in patients with a South African ancestry.

Black South Africans are greatly underrepresented genetic research, the paucity of genomic sequencing data from black South African individuals creates a challenge when cancer screening and risk assessment protocols are designed. The aim of this study was to detect sequence variants predisposing paediatric black South African patients to developing RB.

2.3. The genetic landscape of RB

The first tumor suppressor gene (TSG) described fully was the Retinoblastoma 1 (*RB1*) gene, by Alfred Knudson in 1971, a revolutionary finding that initiated exploration into cancer predisposition genes. More than a decade later, in 1986, the *RB1* gene position was identified by positional cloning (Lee et al., 1987). The explanation of the two-hit hypothesis pioneered our understanding of the molecular basis of cancer (Knudson, 1971). A genetic predisposition to RB malignancy is inherited if a single nonfunctional copy of *RB1* is inherited or present in the germline, a second hit in somatic tissue of the retina will then result in complete loss of function leading to uncontrolled cellular proliferation (Knudson, 1971).

The *RB1* gene exhibits allelic heterogeneity, with more than 1700 different pathogenic germline *RB1* variants logged on the Leiden Open Variation Database (LOVD) (Fokkema et al., 2011). An array of different types of germline variants in *RB1* can serve as an inherited predisposition, or “first hit” (Stenfelt et al., 2017).

Sequence variants such as single nucleotide variants (SNV) leading to a loss-of function in the *RB1* gene product as well as structural variants like whole gene deletions, copy number variation (CNV) (Xu et al., 2020) and rearrangements can render a germline copy of *RB1* non-functional (Davies et al., 2021). Chromothripsis spanning the *RB1* locus provides an additional mechanism of inactivating *RB1* (McEvoy et al., 2014).

With the emergence of sequencing technologies and increasing use thereof, two genetic anomalies with regards to RB was encountered that contradicted the *RB1* two-hit hypothesis. Firstly, in some cases the presence of a pathogenic germline variant in *RB1* is not sufficient to increase the risk to

develop RB or a retinoma neoplasm, as numerous families have a strong family history of a pathogenic variant in *RB1* without the development or RB in successive generations. A nonfunctional *RB1* copy confers genomic instability but additional genetic changes beyond *RB1* may be required for the development of RB, the number of genetic modifications required is greatly debated (Dimaras et al., 2008; Corson & Gallie, 2007). Secondly, RB has been diagnosed in cases with a family history of the disease, indicative of a pathogenic RB1 variant according to the two-hit hypothesis, yet germline sequencing revealed two wild-type copies of *RB1* (Rushlow et al., 2013), suggesting that germline mutations in other tumor predisposition genes beyond *RB1* can act as putative drivers in RB malignancy (Akdeniz et al., 2019).

2.4. Locus heterogeneity observed in RB predisposition

Genes other than RB1 have been linked to heritable RB aetiology. The notion of looking beyond RB1 as provided evidence to add to our knowledge of the genomic landscape of heritable RB. Table 1 summarizes cancer susceptibility genes that have been implicated in RB aetiology, paediatric cancers as well as retinal development pathways.

Copy number analysis pipelines have been designed to detect unbalanced copy number variations, such as large deletions and duplications. In future the addition of copy number analysis should be performed in an aim to detect loss of function variants due to copy number variation in the cancer susceptibility genes mentioned in table 1 (Kooi et al., 2016), bridging the gap between sequence variants and larger structural variants.

Changes at an epigenetic level have been proven to contribute to RB development, with promotor methylation of TSG beyond *RB1* greatly influencing disease progression (Benavente & Dyer, 2015; McEvoy & Dyer, 2015). Non-coding RNAs and miRNAs have been functionally linked to RB tumorigenesis (Reis et al., 2012), creating microRNA (miRNA) signatures and multiomics techniques to be used in clinical settings as both diagnostic and prognostic markers (Delsin et al., 2019; Plousiou & Vannini, 2019).

Germline loss of function mutations in the TSG *p53* has been associated with multiple cancer predisposition syndromes and RB neoplasm development (Kato et al., 1996). Due to the important role *p53* plays in regulating apoptosis in RB (Nork et al., 1997), loss of heterozygosity structural variants involving tumor protein 53 (*p53)* have been associated with RB (Kato et al., 1996). A study conducted

by Castéra, and colleagues discovered a significant contribution of E3 ubiquitin-protein ligase (*MDM2)* as a modifier gene in RB (Castéra et al., 2010) as RB progression is marked by *MDM2* signaling pathway aberrations (Xu et at., 2009). *MDM2* in conjunction with P53-Binding Protein *(MDM4)* regulates the *p53 signaling pathway*. *MDM4* variants implicated in RB oncogenesis were found in a Brazilian cohort (de Oliveira Reis et al., 2012).

The use of exome gene panels detected pathogenic variants in a heritable RB cohort from Istanbul with wild type *RB1* linking RB oncogenesis with the retinoic acid pathway. Pathogenic variants were found in Dehydrogenase [quinone] 1 (*NQO1),* Fibroblast growth factor receptor 4 *(FGFR4),* C-Type Lectin Domain Containing 7A (*CLEC7),* Apolipoprotein C3 *(APOC3),* MutY DNA Glycosylase *(MUTYH),* UDP-Glucuronosyltransferase 1-A *(UGT1A1)*, and RB disease segregated with a variant in Keratin 85 (*KRT85)* (Akdeniz et al., 2019).

The downregulation of Cyclin Dependent Kinase Inhibitor 1A (*CDK4I)*, also rereferred to as *p16INK4a*, is strongly associated with heritable RB but not sporadic RB (Indovina et al., 2009). Germline mutations in *CDKN1A* are linked to uncontrolled cellular proliferation, increasing the risk of developing a neoplasm (Carvalho et al., 2013). Pathogenic variants in oncogenes BHLH Transcription factor (*MYCN)* and Tuberous Sclerosis Complex 2 (*TSC2)* have also been proven to act as major drivers in RB oncogenesis (Francis et al., 2021).

Table 1: Genes implicated in the aetiology of retinoblastoma development

Gene symbol	Full gene name	HGVS Reference transcript	Literature reference	Gene symbol	Full gene name	HGVS Reference transcript	Literature reference
RB1	Retinoblastoma 1	NM_000321.3	Knudson, 1971	*MCCC2*	Methylcrotonyl-CoA Carboxylase Subunit 2	NM_022132.4	Akdeniz et al., 2019
ACADS	Short-chain acyl-CoA dehydrogenase	NM_000017.3	Akdeniz et al., 2019	*MDM2*	E3 ubiquitin-protein ligase	NM_002392.6	Epistolato et al., 2011
APC	APC Regulator Of WNT Signalling Pathway	NM_000038.6	Capasso et al., 2020	*MDM4*	P53-Binding Protein	NM_001204172	de Oliveira Reis et al., 2012
APOC3	Apolipoprotein C3	NM_000040.2	Akdeniz et al., 2019	*MLH1*	MutL Homolog 1	NM_001258271.2	Capasso et al., 2020
BRCA1	BRCA1 DNA Repair associated	NM_007294.4	Capasso et al., 2020	*MSH2*	MutS Homolog 2	NM_000251.3	Capasso et al., 2020
BRCA2	BRCA2 DNA Repair associated	NM_000059.4	Capasso et al., 2020	*MSH3*	MutS Homolog 3	NM_002439.5	Capasso et al., 2020
C2	Complement 2	NM_000063.4	Akdeniz et al., 2019	*MUTYH*[a]	MutY DNA Glycosylase	NM_001128425.1	Akdeniz et al., 2019
CFB	Complement Factor B	NM_001710.5	Akdeniz et al., 2019	*MYCN*	BHLH Transcription factor	NM_005378.4	Francis et al., 2021
CLEC7A[a]	C-Type Lectin Domain Containing 7A	NM_197947.2	Akdeniz et al., 2019	*NQO1*	Dehydrogenase [quinone] 1	NM_000903.2	Akdeniz et al., 2019
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A	NM_000389.5	Carvalho et al., 2013	*PAX6*	Aniridia type II	NM_001368894.2	Li et al., 2011 Meng et al., 2014
CREBBP	CREB Binding Protein	NM_001079846.1	Kooi et al., 2016	*PTCH1*	Patched 1	NM_001083606.3	Capasso et al., 2020
CX3CR1	C-X3-C Motif Chemokine receptor 1	NM_001171174.1	Akdeniz et al., 2019	*p16INK4A*	Cyclin-dependent kinase inhibitor 2A (CDKN2A)	NM_000077.5	Indovina et al., 2009
DSPP[b]	Dentin Sialophosphoprotein	NM_014208.3	Akdeniz et al., 2019	*TP53*	Tumor Protein P53	NM_000546.6	Epistolato et al., 2011
FGFR4	Fibroblast growth factor receptor 4	NM_002011.4	Akdeniz et al., 2019	*RHAG*	Rh Associated Glycoprotein	NM_000324.2	Akdeniz et al., 2019
FUT6	Fucosyltransferase 6	NM_000150.2	Akdeniz et al., 2019	*RPGRIP1*	RPGR Interacting Protein 1	NM_020366.3	Akdeniz et al., 2019
GBE1	Glycogen branching enzyme	NM_000158.3	Akdeniz et al., 2019	*SAMD11*	Sterile Alpha Motif Domain Containing 11	NM_001385640.1	Kubo et al., 2021
GHRL	Ghrelin and Obestatin prepropeptide	NM_001134944.1	Akdeniz et al., 2019	*SERPINA1*	Alpha1AT	NM_001002235.2	Akdeniz et al., 2019
GNPAT	Glycerone-Phosphate O-Acyltransferase	NM_014236.3	Akdeniz et al., 2019	*SLC34A1*	Solute Carrier Family 34 Member 1	NM_003052.4	Akdeniz et al., 2019
HBD	Haemoglobin Delta Chain	NM_000519.3	Akdeniz et al., 2019	*SMAD5*	SMAD Family Member 5	NM_001001419.3	Ueki & Reh, 2012
HFE	Homeostatic Iron Regulator	NM_000410.3	Akdeniz et al., 2019	*TYR*	Tyrosinase	NM_000372.4	Akdeniz et al., 2019
KRT85	Keratin 85	NM_002283.3	Akdeniz et al., 2019	*TSC2*	Tuberous Sclerosis Complex 2	NM_001114382.3	Francis et al., 2021
MBL2[b]	Mannose Binding Lectin 2	NM_000242.2	Akdeniz et al., 2019	*UGT1A1*[a]	UDP-Glucuronosyltransferase 1-A	NM_000463.2	Akdeniz et al., 2019

3. Materials and Methods

3.1. Ethical considerations and consent

This study was conducted as a sub-study under the parent study, Paediatric Cancer (PeCan) Project, Inherited cancer predisposing mutations in a cohort of childhood cancer patients. Ethical clearance was obtained from the University of the Witwatersrand Faculty of Health Sciences Human Research Ethics Committee (Medical) (HREC Medical) prior to the commencement of the research. Ethical clearance certificates for both the parent study, protocol number M180855, and this sub-study, protocol number M230782 are included in appendices A and B respectively. The parents or legal guardians of the participants provided written and/or verbal consent and assent from a subset of participants; copies of the informed consent documents employed in the study are included in appendix F.

3.2. Clinical presentation and patient demographics

Eight paediatric patients diagnosed with RB were included in this study from the PeCan project. All eight patients are of black South-African ancestry. No family history of RB was reported in any of the patients. Table 2 summarizes the clinical features and demographics of the eight patients.

Table 2: Clinical presentation of RB patients

Patient ID [a]	**Sex** [b]	**Age of diagnosis**	**Laterality of RB**
PC26	Female	1 Year 8 Months	Unilateral RB
PC43	Female	1 Year 6 Months	Unilateral RB
PC50	Female	4 Years	Unilateral RB
PC79	Male	4 Months	Unilateral RB
PC90	Female	3 Year 6 Months	Unilateral RB
PC104	Female	8 Months	Bilateral RB
PC120	Male	3 Months	Bilateral RB
PC123	Male	9 Months	Unilateral RB

[a] Patient IDs used to anonymize patient information.

[b] Sex as assigned at birth

3.3. Whole exome sequencing

3.3.1. DNA sample preparation and sample quality control

Genomic DNA was extracted and purified from peripheral blood samples prior to the commencement of this project using the salt-out method (Maurya et al., 2013). DNA quantification commenced with assessing the concentration of double-stranded genomic DNA using the Qubit ™ High Sensitivity assay and the Qubit ™ fluorometer (ThermoFisher Scientific, Waltham MA, USA). Concentration of all samples were normalised to 10 ng/µl.

3.3.2. Library generation

Target amplification was performed using Ion AmpliSeq ™ Exome RDY reagents, followed by partial digestion of amplicons using the Ion AmpliSeq ™ FuPa reagent. Ligation of Ion P1 Adapters and Ion Xpress ™ Barcodes was completed as per ThermoFisher Scientific manufacturer's guidelines (ThermoFisher Scientific, Waltham MA, USA). Barcoded libraries were purified using Agencourt ™ AMPure XP Reagents (Beckman Coulter, Brea, CA, USA) followed by library amplification. Amplified libraries were purified in two rounds, again using Agencourt ™ AMPure XP Reagents (Beckman Coulter, Brea, CA, USA).

Libraries were quantified using the 4200 TapeStation ™ (Agilent technologies, Santa Clara, USA) in conjunction with the prescribed High Sensitivity D5000 ScreenTape assay. The concentration of 200-400 base pair fragments were measured for the purposes of quality control and diluted accordingly to facilitate the combination of exome libraries.

3.3.3. Templating and sequencing

Two Ion AmpliSeq ™ Exome RDY libraries were combined on a single Ion 540 ™ Chip, templated in the Ion 540 ™ Kit – Chef (ThermoFisher Scientific, Waltham MA, USA) as planned on Torrent Suite™ software. Sequencing was performed using the Ion Torrent GeneStudio ™ S5 Sequencer (ThermoFisher Scientific, Waltham MA, USA). Sequencing output included variant call files (VCF), binary alignment maps (BAM), index files and quality control reports.

3.4. Variant interpretation

VCF files produced were annotated with Ensembl variant effect predictor (VEP) tool (GRCh37 version 110) (McLaren et al., 2016) available online from Ensembl 110 Genome Browser (https://www.ensembl.org) (Martin et al., 2023). Note that the decision was made to align all exomes to the human reference genome build 19 (GRCh37/Hg19), for the sake of consistency throughout the bioinformatics workflow. Annotation outputs included variant location and variant type, i.e., single nucleotide variants (SNVs) such as missense or nonsense variants, insertions, or deletions (indels), frameshift, or splice variant (Hunt et al., 2022). Annotation output also included variant allele frequency as reported by two genomic databases, 1000 genomes project (The 1000 Genomes Project Consortium et al., 2015) and the genome Aggregation Database (gnomAD) (Chen et al., 2022). All the above-mentioned databases were accessed between August and September 2023.

First line of variant filtering was to prioritize the genes that have previously been reported to be implicated in RB predisposition and paediatric cancer predisposition. These genes are summarized in Table 1. Variants with a minor allele frequency (MAF) above 5% were filtered out. Lastly, variants that were predicted to have little to no impact on protein structure and/ or function were filtered out. The in-silico tools used to quantify variant impact on protein structure and function included Sort Intolerable from Tolerable (SIFT) (Ng & Henikoff, 2003), Polymorphism Phenotyping version 2 (POLYPHEN-2) (Adzhubei et al., 2010). Combined Annotation-Dependent Depletion (CADD) scores were allocated to SNVs and small indels based on how deleterious that variant is predicted to be (Rentzsch et al., 2019). Both SIFT and POLYPHEN-2 take the variant position's degree of evolutionary conservation into account as part of predicting the impact of the variant on a protein level. These in silico tools predict what structural and functional changes may occur in a protein due to amino acid substitution (de la Campa et al. 2017). Lastly, variants of interest were modelled on MutationTaster2021 to distinguish polymorphisms from mutations (Steinhaus et al., 2021).

Coverage and the quality of variant calling was assessed for the shortlisted variants by viewing binary alignment maps (BAM) and their index files in an integrative genomics viewer (IGV) (Robinson et al., 2011), variants with 30x coverage or greater were considered to be true variant calls. Variants were classified according to The American College of Medical Genetics and Genomics and Association for Molecular Pathology (ACMG/AMP) guidelines (Richards et al., 2015). Previously reported variants were analysed based on ClinVar submissions (Landrum et al., 2018).

4. Results

4.1. Clinical presentation of patients

The eight patients had a wide distribution of age of onset, ranging from as early as 3 months to as late as 4 years. Mean age of onset for RB cohort was 19 months with a standard deviation of 17.22 months. Both male (37.5%) and female (62.5%) patients were included in the cohort. The two patients (25%) with the earliest age of onset (3 months and 8 months respectively) were the only patients that presented with bilateral RB, while the remaining six patients (75%) presented with unilateral RB (onset age range ranging from 4 months to 48 months).

4.2. Variant annotation

Whole exome sequencing (WES) of eight paediatric RB patient with a black South African ancestry collectively returned 353 655 annotated variants spanning the entire genome. Figure 1 depicts the distribution of variant consequences as observed in the 8 patient cohort. The largest number of variants were 122 011 intronic variants, accounting for 34.5%. These intronic variants were filtered out immediately. The second highest proportion of variants were 54 392 synonymous variants (15.38%), followed by 38 018 missense variants (10.75%). The 25 640 downstream variants made up 7.25% of the total number of variants. A total of 21 679 variants (6.13%) were detected in regulatory regions, and 18 143 variants (5.13%) in upstream regions. Collectively, non-coding transcripts, non-coding transcript exon variants and splice site polymorphisms contributed 73 772 variants (20.86%), grouped as other in Figure 1. The variants in coding regions were mainly synonymous and missense variants as depicted in Figure 2. A total of 4 538 novel variants were annotated.

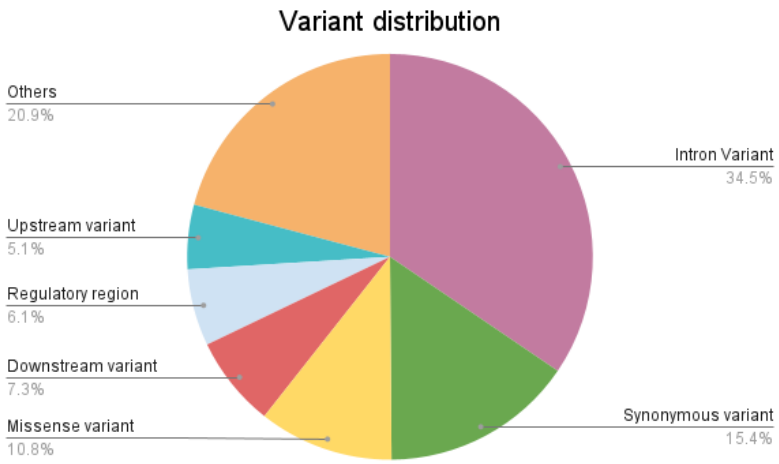


Figure 1: Distribution of annotated variants

The majority of the sequence variants annotated were intronic variants, 122 011 intronic variants accounting for 34.5% of all variants, followed by 54 392 synonymous variants (15.38%). Thus approximately half (49.9%) of the annotated variants were filtered out at a very early stage of variant prioritization.

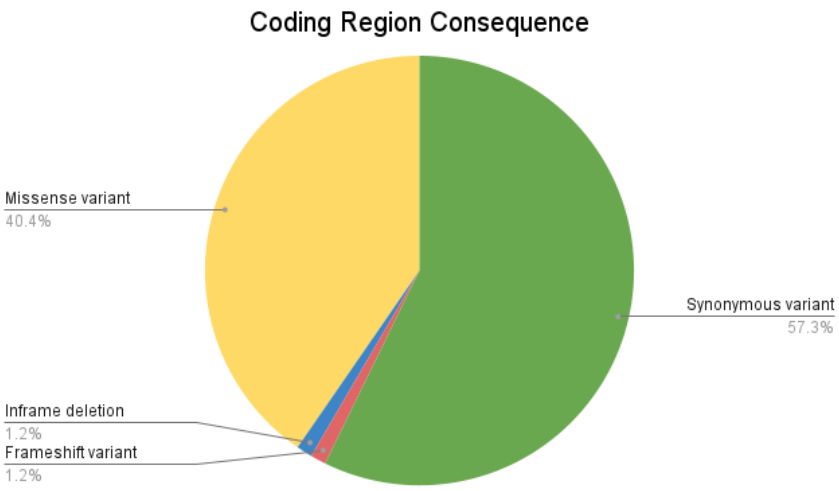


Figure 2: Distribution of variants in coding regions

The vast majority of annotated variants in the coding region predicted to have an impact on protein structure and function were 38 018 missense variants

4.3. *RB1* sequence variants

All exonic variants in *RB1* are tabulated in appendix C. Two benign *RB1* sequence variants, NM_000321.3:c.1574C>G (NP_000312.2:p.Ala525Gly) and NM_000321.3:c.45_53del (NP_000312.2:p.Ala16_Ala18del) were detected in a single patient, PC43. The beforementioned in-frame deletion of three amino acids NM_000321.3:c.45_53del (NP_000312.2:p.Ala16_Ala18del) was shared in three of the eight patients (PC43, PC50 and PC123). A single missense variant NM_000321.3:c.1574C>G (NP_000312.2:p.Ala525Gly) was found in a single patient (PC26).

4.4. Variants of uncertain significance

Variants of interest in each patient classified as variants of uncertain significance are summarized in Table 3, along with the ACMG/AMP codes applied to each variant. All of the variants are classified as variants of uncertain significance (VUS) because the ACMG/AMP codes applied could not categorize the variants as likely benign or likely pathogenic. Variants in blue are shared between individuals in the cohort and are discussed separately. An extended list of variants in cancer predisposing genes associated with RB, MAF < 5% are tabulated in appendix D.

A total of 12 frameshift variants predicted to have a detrimental impact on protein structure and protein folding were detected. All patients carried as least one frameshift variant. Two novel frameshift mutations were classified as VUS. A novel frameshift variant in *PTCH1* NM_001385640.1:c.2386C>T (NP_001372569.1:p.Leu796Phe) resulting in a truncated protein was detected in patient 26 and novel frameshift mutation in *PTCHD3* NM_001034842.4:c.152del, (NP_001030014.2:p.Pro51ArgfsTer69) detected in patient 120 diagnosed at a young age with bilateral tumor presentation.

Two previously reported frameshift variants were detected in a single patient only. Frameshift variant in FUT6, rs150560931, caused by a single nucleotide insertion NM_000150.3:c.501_502insC (NP_000141.1:p.Tyr168LeufsTer230) and a frameshift variant in *CREBBP,* rs587783507, NM_001079846.1:c.5723del (NP_001073315.1:p.Pro1908HisfsTer30). The TSG *MSH3 contained a novel inframe deletion,* NM_002439.5:c.162_179del NP_002430.3:p.Ala57_Ala62del and a previously reported missense variant rs34168832 NM_002439.5:c.1817G>A (NP_002430.3:p.Ser606Asn) in two respective patients.

The vast majority of variants were missense variants, two of which were novel. The novel missense variant were detected in *APC* NM_000038.6:c.1251T>G (NP_000029.2:p.Cys417Trp) and *MBL2* NM_000242.3:c.34_35delinsGG (NP_000233.1:p.Leu12Gly). Three previously reported missense variants were predicted to be damaging by more than one in-silico prediction tool. DNA mismatch repair genes, *MSH3,* rs34168832, NM_002439.5:c.1817G>A (NP_002430.3:p.Ser606Asn) and *MLH1* rs63751244 NM_001167617.3:c.1375G>A (NP_001161089.1:p.Glu459Lys). *UGT1A1,* NM_000463.3:c.22G>A (NP_000454.1:p.Gly8Arg) was also deemed a VUS.

For patient PC43 diagnosed at 18 months, presenting with unilateral RB, a total of 14 variants in cancer predisposing genes were annotated, the greatest number of variants per patient. Twelve out of the fourteen variants were missense variants predicted to have a negligible impact on protein function. By factoring in Clinvar record for the existing missense variants, all were classified as likely benign based on codes PM2 and BP4. The frameshift variant *SAMD11* NM_001385640.1:c.1890_1891del (NP_001372569.1:p.Lys631ArgfsTer71) variant in patient PC43 is shared between individuals in the cohort and deemed a VUS. Patient PC123 also carried the frameshift variant in *SAMD11* NM_001385640.1:c.1890_1891del (NP_001372569.1:p.Lys631ArgfsTer71), along with a missense variant in *HFE* rs28934889. Table 4 contains the ACMG/AMP codes applied to shared variants.

Table 3: Subset of variants of interest as observed per patient

Gene	Variant HGVSc and Variant HGVSp	rsID	Consequence	Impact [a]	ACMG/AMP [b]
PC 26					
PTCH1	NM_000264.5:c.3913del, NP_000255.2:p.Asp1305ThrfsTer67	novel	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
MLH1	NM_001167617.3:c.1375G>A, NP_001161089.1:p.Glu459Lys	rs63751244	missense variant	MODERATE	VUS (PM2, PP3)
UGT1A1	NM_000463.3:c.22G>A, NP_000454.1:p.Gly8Arg	rs749552053	missense variant	MODERATE	VUS (PM2, PP3)
PC 43					
SAMD11	NM_001385640.1:c.1890_1891del, NP_001372569.1:p.Lys631ArgfsTer71	novel	frameshift variant	HIGH	VUS (PM2, PP3,PM4)
PC 50					
APC	NM_000038.6:c.1251T>G, NP_000029.2:p.Cys417Trp	-	missense variant	MODERATE	VUS (PM2, PP3)
SMAD5	NM_001001419.3:c.1313-1_1313insC, NP_001001419.1:p.Asn438ThrfsTer11	-	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
CREBBP	NM_001079846.1:c.5723del, NP_001073315.1:p.Pro1908HisfsTer30	rs587783507	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
PC 79					
SMAD5	NM_001001419.3:c.1313-1_1313insC, NP_001001419.1:p.Asn438ThrfsTer11	-	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
MSH 3	NM_002439.5:c.1817G>A, NP_002430.3:p.Ser606Asn	rs34168832	missense variant	MODERATE	VUS (PM2, PP3)
PC 90					
SAMD11	NM_001385640.1:c.1890_1891del, NP_001372569.1:p.Lys631ArgfsTer71	novel	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
SMAD5	NM_001001419.3:c.1313-1_1313insC, NP_001001419.1:p.Asn438ThrfsTer11	-	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
PC 104					
SAMD11	NM_001385640.1:c.1890_1891del, NP_001372569.1:p.Lys631ArgfsTer71	novel	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
PC 120					
PTCHD3	NM_001034842.4:c.152del, NP_001030014.2:p.Pro51ArgfsTer69	-	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
SMAD5	NM_001001419.3:c.1313-1_1313insC, NP_001001419.1:p.Asn438ThrfsTer11	-	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
PC 123					
SAMD11	NM_001385640.1:c.1890_1891del, NP_001372569.1:p.Lys631ArgfsTer71	novel	frameshift variant	HIGH	VUS (PM2, PP3, PM4)
HFE	NM_139006.3:c.157G>A, NP_620575.1:p.Val53Met	rs28934889	missense variant	MODERATE	VUS (PM2, PP3)

[a] The impact of the variant of protein structure and function is a prediction based on scored collected from multiple in silico prediction tools, including SIFT, POLYPHEN-2 and CADD.

[b] Codes in brackets applied according to ACMG/AMP guidelines to classify variant

-: Previously reported variants without an rsID at the time of interpretation

4.5. Shared variants

Several of the genes proposed as candidate genes in RB or paediatric cancer, as mentioned in Table 1 returned variants observed in multiple patients despite the low allele frequency in the general population. Table 4 depicts the shared variants, the frequency at which the mutation occurred in the eight patients in the study cohort and the allele frequency in the general population. Two frameshift variants, *SAMD11* NM_001385640.1:c.1890_1891del (NP_001372569.1:p.Lys631ArgfsTer71) and *SMAD5* NM_001001419.3:c.1313-1_1313insC (NP_001001419.1:p.Asn438ThrfsTer11) were found in half the patients. Coverage and quality of variant calling was evaluated by visualizing BAM files on an IGV. The IGV images of a select few variants are depicted in appendix E to demonstrate the BAM images utilized to assess variant calling quality.

Table 4: Shared variants (SV)

Gene	**Variant HGVSc notation and Variant HGVSp notation**	**Variant frequency [a]**	**MAF [b]**	**rsID**	**ACMG/AMP Classification [c]**
HFE	NM_139006.3:c.157G>A NP_620575.1:p.Val53Met	2 (25%)	0.0002	rs28934889	VUS (PM2, PP3)
KRT85	NM_002283.4:c.233G>A NP_002274.1:p.Arg78His	3 (37.5%)	0.0328	rs61630004	Likely benign (PM2, PP1, BP6)
SAMD11	**NM_001385640.1:c.1890_1891del NP_001372569.1:p.Lys631ArgfsTer71**	4 (50%)	-	Novel variant	VUS (PM2, PP3, PM4)
SMAD5	**NM_001001419.3:c.1313-1_1313insC NP_001001419.1:p.Asn438ThrfsTer11**	4 (50%)	-	Novel variant	VUS (PM2, PP3, PM4)

[a] Variant frequency in cohort of eight RB patients

[b] Minor allele frequency in general population from gnomAD

[c] Codes in brackets applied according to ACMG/AMP guidelines to classify variant

Variants in bold are novel variants predicted to have a deleterious impact on protein structure and function according to in-silico prediction tools, including SIFT, POLYPHEN-2 and CADD.

The majority of variants classified as VUS were missense variants, however the variants predicted to have the largest impact on protein structure and function were the SAMD11 and SMAD5 frameshift variants shared between individuals. Both of the frameshift variants mentioned are novel variants occurring in 50% of the RB population of the study.

5. Discussion

The germline genetic contribution to RB aetiology has evolved beyond *RB1,* but little to no investigation has probed into the genetic landscape of RB in South African patients of African ancestry. This study aimed to address this disparity and gap in genomic knowledge by testing a cohort of eight paediatric RB patients with a South African ancestry for germline sequence variants predisposing them to RB malignancy. WES sequencing data uncovered 22 variants in cancer predisposition genes with an uncertain significance and 4 538 novel variants distributed across the genome. The lack of functional and segregation data for the novel variants in cancer predisposition genes reduced them to a VUS classification.

The correlation of the young age of onset with bilateral tumor presentation is characteristic of individuals carrying a germline mutation or "first hit" (Knudson, 1971). This correlation was observed in the cohort, patients PC104 and PC120 both presented with bilateral RB at a young age, exemplifying that the two hit hypothesis still holds true, the presence of a constitutional germline mutation in cancer susceptibility gene can be observed as bilateral RB presentation at a young age. No clear pathogenic or likely pathogenic sequence variant was observed in *RB1* in patients PC104 or PC120. In approximately half of the patients, no exonic sequence variant impacting protein structure was observed in *RB1*, as presented in appendix 3. The absence of likely pathogenic or pathogenic sequence variants in RB1 describe the RB phenotype cements the notion that variants in cancer susceptibility genes beyond *RB1* need to be investigated to determine the full germline mutational spectrum associated with heritable RB. Patient 104, who was diagnosed at 8 months with bilateral RB had a mere two sequence variants of interest in cancer predisposing genes, neither of which explain the RB phenotype.

5.1. *RB1* variants

The in-frame deletion NM_000321.3:c.45_53del NP_000312.2:p.Ala16_Ala18del present in three patients, has previously been described (rs572454921) and is predicted to be benign because it result in a negligible change in protein structure. One patient, PC 26 had a *RB1* missense variant rs4151539 (NM_000321.3:c.1574C>G, NP_000312.2:p.Ala525Gly) previously described in the literature as probably pathogenic, in a Brazilian RB cohort that screened for *RB1* mutations in both RB and retinoma patients (Barbosa et al., 2013). As the reporting of variants observed in sequencing data became good scientific practice, conflicting interpretations of this variant emerged. This variant is classified as likely benign.

5.2. Shared variants

Variants shared between the patients in the RB cohort at a higher allele frequency than the allele frequency observed in the general population are of interest. The majority of these shared variants are found in genes associated with retinal metabolism and rod photoreceptor development (Gao et al., 2009). The variants observed are yet to be linked to RB with functional studies. The shared variants are depicted in blue and summarized in Table 3, with the ACMG/AMP codes applied.

A VUS in the Homeostatic Iron Regulator (*HFE)* gene was detected in two individuals, previously identified as rs28934889, NM_139006.3:c.157G>A, NP_620575.1:p.Val53Met. Akdeniz and colleagues were the first to propose HFE as a candidate gene in RB (Akdeniz et al., 2019). The *HFE* gene is responsible for iron metabolism in the retina, optimal iron metabolism and regulation prevents oxidative damage due to iron overload (Gnana-Prakasam et al., 2010). The HFE complexed with other proteins regulate iron concentrations in the electron transport chain through the SMAD1/5/8 pathway (Gao et al., 2009).

The *KRT85* rs61630004 (NM_002283.4:c.233G>A NP_002274.1:p.Arg78His) missense mutation observed in three of the eight RB patients has been described in literature as a candidate gene in RB etiology (Akdeniz et al., 2019), suggesting the variant might segregate with disease, leading to the application of ACMG/AMP code *PP1*. A study conducted in Istanbul observed this variant in multiple individuals affected with RB, two of them in the same family, implying that the variant might segregate with disease. Due to the significantly higher frequency of variant in disease cohort compared to the general population ACMG/AMP code PM2 was applied. The prediction of a deleterious effect by an in-silico tool the variant was classified as pathogenic by Akdeniz and colleagues.

The in-silico predictors used has since become outdated, improved versions of the same in-silico tools, SIFT (Ng & Henikoff, 2003) and POLYPHEN-2 (Adzhubei et al., 2010) now predict rs61630004 (NM_002283.4:c.233G>A NP_002274.1:p.Arg78His) to have inconsequential effects on protein structure and function. MutationTaster2021 (Steinhaus et al., 2021) classified this variant as a polymorphism as opposed to a damaging mutation. The application of ACMG/AMP codes *PM2, PP1 and BP6* collectively lead to the current classification of *KRT85* NM_002283.4:c.233G>A NP_002274.1:p.Arg78His as likely benign which is in accordance with Clinvar classification.

The notably high frequency (50%) of a novel frameshift variant in *SAMD11* NM_001385640.1:c.1890_1891del brings the impact of this *SAMD11* on RB predisposition in question. The absence of this variant in population databases prompted the application of the ACMG/AMP code *PM2*. A truncated protein is produced, predicted to have a detrimental effect on protein function, as a result, ACMG/AMP codes *PP3* and *PM4* were

applied. Unfortunately, the genotype-phenotype correlation is difficult to establish. A recent functional study demonstrated that *SAMD11* is involved in establishing rod photoreceptor identity in retinal cells during retinal cell differentiation and development by forming part of the rod-specific polycomb recessive complex (Kubo et al., 2021). Dysregulation of *SAMD11* is one of the numerous genetic causes of X-linked recessive retinitis pigmentosa (Corton et al., 2016).

A second novel frameshift mutation was observed in the *SMAD5,* a gene that similar to *SAMD11* is expressed in retinal tissue. In vivo experiments provide functional proof that *SMAD5 i*s activated to repair retinal cells in response to light damage. Identical to the *SAMD11* variant discussed above, the novel nature of the *SMAD5* variant prompted the application of the *PM2* code. Furthermore, the *PP3* code was applied because in-silico tools predicted this frameshift variant to be deleterious. *SMAD5* is a component of the BMP-SMAD1/5/8 neuroprotective signaling pathway promote survival of retinal ganglion cells (Ueki & Reh, 2012). In half of the patients in the RB cohort, a germline NM_001001419.3:c.1313-1_1313insC variant in *SMAD5* was observed, predicted to cause a lack of function. Future functional studies are needed to speak to whether the inability to repair retinal cells damaged by light can create genomic instability, increasing the likelihood of future mutations and clonal evolution. Both patients PC104 and PC120 presented with an early age of onset and bilateral tumours, suggestive of hereditary RB. The lack of sequence variants capable of rationalizing the RB phenotype shines a light on the limitation of analyzing sequence variants exclusively.

5.3. Variants of uncertain significance

The frameshift mutation in CREB Binding Protein (*CREBBP)* NM_001079846.1:c.5723del (NP_001073315.1:p.Pro1908HisfsTer30), produces a truncated protein product. *CREBBP* has been established as a TSG in lymphoma (Zhang et al., 2017) with haploinsufficient function. Loss of function of *CREBBP* is the autosomal dominant genetic cause of Rubinstein-Taybi syndrome, a neurodevelopmental disorder associated with a predisposition to developing tumours of the nervous system (Roelfsema & Peters, 2007). Somatic *CREBBP* loss of function mutations and downregulation have been proven to be putative drivers in RB tumorigenesis (Kooi et al., 2016), however the impact of germline deleterious mutations in *CREBBP* on RB aetiology remains unknown, this variant has not been reported in combination with RB.

PC 50, diagnosed with unilateral RB at the age of 4 years, presented with a germline variant in the Adenomatous polyposis coli gene (*APC)* NM_000038.6:c.1251T>G, NP_000029.2:p.Cys417Trp, TSG inactivated in a large proportion of colorectal cancers (Sparks et al., 1998), a substructure of *APC* called *APC2* has been established as a TSG in RB (Singh et al., 2016; Beta et al., 2015).Wild-type *APC* antagonizes the WNT signaling pathway, to negatively regulate cell growth. Mutations in genes regulating the WNT signaling

pathway have been linked to vitreoretinopathies like retinopathy of prematurity (Drenser, 2016). Case studies have been reported where patients with regressed retinopathy developed RB after approximately 4 years, but due to the small sample size, a causative link between retinopathy and RB could not be established with adequate statistical power (Vasu et al., 2011). The late age of onset and unilateral RB presentation in PC 50 are indicative of sporadic RB, the absence of a driver mutation is reasonable.

From the beforementioned study by Singh and colleagues, defective mismatch repair was implicated in RB (Singh et al., 2016) by increasing genomic instability. Two DNA mismatch repair genes, *MLH1* and *MSH3*, observed in two patients, PC26 and PC79 respectively, returned variants deemed as variants of uncertain significance.

Patient PC26 presented with a frameshift mutation in Patched gene 1 (*PTCH1*), predicted to be deleterious by in-silico prediction tools (ACMG/AMP code *PP3* applied). *PTCH1* is a TSG implicated in central nervous system cancers such as medulloblastoma (Skowron et al., 2021) and carcinoma due to dysregulation of the sonic hedgehog signaling pathway leading to inappropriate cellular proliferation (Doheny et al., 2020).

The patient predicted to be genetically predisposed to RB malignancy due to early age of onset and bilateral tumours, patient (PC120) presented with a novel frameshift variant in *PTCHD3* NM_001034842.4:c.152del, NP_001030014.2:p.Pro51ArgfsTer69 where a truncated protein is produced. When investigating the germline mutational spectrum of predisposing genes implicated in colorectal cancer, *PTCHD3* was proposed as a candidate predisposition gene that may serve as a TSG (Smith et al., 2013).

5.4. Future prospects and limitations

The majority of final variant classifications were VUS. In the case that functional data and segregation data becomes available for a variant classified as VUS, the variant will be revisited and reclassified in accordance with ACMG/AMP guidelines. The inability to find previously reported driver mutation in this cohort, especially in the patients with a young age of onset and bilateral RB, brings the utility of analyzing only sequence variants from WES data in question. This study had several limitations, mainly the failure of WES to detect large balanced structural anomalies and large deletions and duplications.

The addition of copy number variation assays such as multiplex ligation-dependent probe amplification (MLPA) (Livide et al., 2012) or chromosomal microarrays (CMA) can complement WES sequencing data, producing a more comprehensive picture of all germline variants present in an individual. Epigenetic changes such as promotor methylation (McEvoy & Dyer, 2015) and non-coding RNA molecules (Plousiou & Vannini, 2019) that contribute to RB etiology went undetected when sequencing on a genomic level.

The underrepresentation of African populations in cancer genomics cohorts inevitably complicates variant interpretation, exemplified by the numerous novel variants found in this study. Cancer is an extremely heterogenous disease. The possibility of patients carrying a germline variant in genes not included in the curated list (Table 1) cannot be discounted. The study cohort included patients of both sexes, where the age of onset varied widely, and RB presentation was unilateral in some and bilateral in others. As the majority of RB cases are sporadic, it is reasonable to assume that some patients will not carry a predisposing constitutive variant.

6. Conclusion

The utmost degree of sequence variant classification according to the ACMG/AMP guidelines was variant of unknown significance. The inability to link the RB phenotype of the testing cohort to a germline sequence variant in the genes previously associated with RB supports the beyond *RB1* hypothesis. Although *RB1* is likely to remain the most likely driver in RB etiology, a wider testing strategy including CNV will yield more informative results. Understanding the full mutational landscape of RB will aid in the construction of appropriate treatment plans for RB patients and genetic counselling for family members.

Variant classification in understudied and underrepresented populations, like the black South African population is hampered by the lack of information available. The genomic data currently available is a far cry from being representative of a genetically diverse country like South Africa. This study adds to the ever-growing collection of variants found in the black South African population, a critical set piece to expanding the genomic research resources in South Africa.

Competing interests

The authors declare that they have no competing interests.

Statement on data availability

Data is not available in public domain at present. All variants will be made available on Clinvar https://www.ncbi.nlm.nih.gov/clinvar/ .

Funding

This study was funded by the Inherited cancer predisposing mutations in a cohort of childhood cancer patients study, principal investigator Dr Lindie Lamola.

7. References

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038%2Fnmeth0410-248

Akdeniz, D., Tuncer, S. B., Kebudi, R., Celik, B., Kuru, G., Kilic, S., Sukruoglu Erdogan, O., Avsar, M., Buyukkapu Bay, S., Tuncer, S., & Yazici, H. (2019). Investigation of new candidate genes in retinoblastoma using the TruSight One "clinical exome" gene panel. *Molecular Genetics & Genomic Medicine*, *7*(8), 1-14. https://doi.org/10.1002/mgg3.785

Barbosa, R. H., Aguiar, F. C., Silva, M. F., Costa, R. A., Vargas, F. R., Lucena, E., Carvalho de Souza, M., de Almeida, L. M., Bittar, C., Ashton Prolla, P., Bonvicino, C. R., & Seuánez, H. N. (2013). Screening of *RB1* alterations in Brazilian patients with retinoblastoma and relatives with retinoma: phenotypic and genotypic associations. *Investigative Ophthalmology & Visual Science*, *54*(5), 3184–3194. https://doi.org/10.1167/iovs.13-11686

Benavente, C. A., & Dyer, M. A. (2015). Genetics and epigenetics of human retinoblastoma. *Annual Review of Pathology*, *10*, 547–562. https://doi.org/10.1146/annurev-pathol-012414-040259

Beta, M., Chitipothu, S., Khetan, V., Biswas, J., & Krishnakumar, S. (2015). Hypermethylation of adenomatosis polyposis coli-2 and its tumor suppressor role in retinoblastoma. *Current eye research*, *40*(7), 719–728. https://doi.org/10.3109/02713683.2014.954673

Bouchoucha, Y., Matet, A., Berger, A., Carcaboso, A. M., Gerrish, A., Moll, A., Jenkinson, H., Ketteler, P., Dorsman, J. C., Chantada, G., Beck-Popovic, M., Munier, F., Aerts, I., Doz, F., Golmard, L., & European Retinoblastoma Group EuRbG (2023). Retinoblastoma: From genes to patient care. *European Journal of Medical Genetics*, *66*(1), 104674. https://doi.org/10.1016/j.ejmg.2022.104674

Capasso, M., Montella, A., Tirelli, M., Maiorino, T., Cantalupo, S., & Iolascon, A. (2020). Genetic Predisposition to Solid Pediatric Cancers. *Frontiers in Oncology*, *10*, 590033. https://doi.org/10.3389/fonc.2020.590033

Carvalho, I. N. S. R., Reis, A. H. de O., Cabello, P. H., & Vargas, F. R. (2013). Polymorphisms of CDKN1A gene and risk of retinoblastoma. *Carcinogenesis*,*34*(12),2774–2777. https://doi.org/10.1093/carcin/bgt308

Castéra, L., Sabbagh, A., Dehainault, C., Michaux, D., Mansuet-Lupo, A., Patillon, B., Lamar,E., Aerts, I., Lumbroso-Le Rouic, L., Couturier, J., Stoppa-Lyonnet, D., Gauthier- Villars, M., & Houdayer, C. (2010). MDM2 as a Modifier Gene in Retinoblastoma. *JNCI:Journal of the National Cancer Institute*, *102*(23), 1805– 1808. https://doi.org/10.1093/jnci/djq416

Chen, S., Francioli, L. C., Goodrich, J. K., Collins, R. L., Wang, Q., Alföldi, J., Watts, N. A., Vittal, C., Gauthier, L. D., Poterba, T., Wilson, M. W., Tarasova, Y., Phu, W., Yohannes, M. T., Koenig, Z., Farjoun, Y., Banks, E., Donnelly, S., Gabriel, S., Gupta, N., Ferriera, S., Tolonen, C., Novod, S., Bergelson, L., Roazen, D., Ruano-Rubio, V., Covarrubias, M., Llanwarne, C., Petrillo, N., Wade, G., Jeandet, T., Munshi, R., Tibbetts, K., gnomAD Project Consortium, O'Donnell-Luria, A., Solomonson, M., Seed, C., Martin, A. R., Talkowski, M. E., Rehm, H. L., Daly, M. J., Tiao, G., Neale, B. M.†, MacArthur, D. G.† & Karczewski, K. J. (2022). A genome-wide mutational constraint map quantified from variation in 76,156 human genomes. *bioRxiv* 2022.03.20.485034 https://doi.org/10.1101/2022.03.20.485034

Corson, T. W., & Gallie, B. L. (2007). One hit, two hits, three hits, more? Genomic changes in the development of retinoblastoma. *Genes, Chromosomes & Cancer*, *46*(7), 617–634. https://doi.org/10.1002/gcc.20457

Corton, M., Avila-Fernández, A., Campello, L., Sánchez, M., Benavides, B., López-Molina, M. I., Fernández-Sánchez, L., Sánchez-Alcudia, R., da Silva, L. R. J., Reyes, N., Martín-Garrido, E., Zurita, O., Fernández-San José, P., Pérez-Carro, R., García-García, F., Dopazo, J., García-Sandoval, B., Cuenca, N., & Ayuso, C. (2016). Identification of the Photoreceptor Transcriptional Co-Repressor SAMD11 as Novel Cause of Autosomal Recessive Retinitis Pigmentosa. *Scientific reports*, *6*, 35370. https://doi.org/10.1038/srep35370

Davies, H. R., Broad, K. D., Onadim, Z., Price, E. A., Zou, X., Sheriff, I., Karaa, E. K., Scheimberg, I., Reddy, M. A., Sagoo, M. S., Ohnuma, S. I., & Nik-Zainal, S. (2021). Whole-Genome Sequencing of Retinoblastoma Reveals the Diversity of Rearrangements Disrupting *RB1* and Uncovers a Treatment-Related Mutational Signature. *Cancers*, *13*(4), 754. https://doi.org/10.3390/cancers13040754

de la Campa, E. álvarez, Padilla, N., & de la Cruz, X. (2017). Development of pathogenicity predictors specific for variants that do not comply with clinical guidelines for the use of computational evidence. BMC Genomics, 18(Suppl 5), 1–14. https://doi.org/10.1186/s12864-017-3914-0

Delsin, L. E. A., Salomao, K. B., Pezuk, J. A., & Brassesco, M. S. (2019). Expression profiles and prognostic value of miRNAs in retinoblastoma. *Journal of cancer research and clinical oncology*, *145*(1), 1–10. https://doi.org/10.1007/s00432-018-2773-7

de Oliveira Reis, A. H., de Carvalho, I. N. S. R., de Sousa Damasceno, P. B., Ferman, S. E.,Lucena, E., Lopez-Camelo, J. S., Seuánez, H. N., & Vargas, F. R. (2012). Influence of MDM2 and MDM4 on development and survival in hereditary retinoblastoma. *Pediatric Blood & Cancer*, *59*(1), 39–43. https://doi.org/10.1002/pbc.24014

Dimaras, H., Khetan, V., Halliday, W., Orlic, M., Prigoda, N. L., Piovesan, B., Marrano, P., Corson, T. W., Eagle, R. C., Jr, Squire, J. A., & Gallie, B. L. (2008). Loss of RB1 induces non-proliferative retinoma: increasing genomic instability correlates with progression to retinoblastoma. *Human Molecular Genetics*, *17*(10), 1363–1372. https://doi.org/10.1093/hmg/ddn024

Doheny, D., Manore, S. G., Wong, G. L., & Lo, H. W. (2020). Hedgehog Signaling and Truncated GLI1 in Cancer. *Cells*, *9*(9), 2114. https://doi.org/10.3390/cells9092114

Drenser K. A. (2016). Wnt signaling pathway in retinal vascularization. *Eye and brain*, *8*, 141–146. https://doi.org/10.2147/EB.S94452

Epistolato, M. C., Disciglio, V., Livide, G., Berchialla, P., Mencarelli, M. A., Marozza, A., Amenduni, M., Hadjistilianou, T., De Francesco, S., Acquaviva, A., Toti, P., Cetta, F., Ariani, F., De Marchi, M., Renieri, A., & Giachino, D. (2011). p53 Arg72Pro and MDM2 309 SNPs in hereditary retinoblastoma. *Journal of Human Genetics*, *56*(9), 685–686. https://doi.org/10.1038/jhg.2011.82

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., & den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. *Human Mutation*, *32*(5), 557–563. https://doi.org/10.1002/humu.21438

Francis, J. H., Richards, A. L., Mandelker, D. L., Berger, M. F., Walsh, M. F., Dunkel, I. J., Donoghue, M. T. A., & Abramson, D. H. (2021). Molecular Changes in Retinoblastoma beyond *RB1*: Findings from Next-Generation Sequencing. *Cancers*, *13*(1), 149. https://doi.org/10.3390/cancers13010149

Gao, J., Chen, J., Kramer, M., Tsukamoto, H., Zhang, A. S., & Enns, C. A. (2009). Interaction of the hereditary hemochromatosis protein HFE with transferrin receptor 2 is required for transferrin-induced hepcidin expression. *Cell metabolism*, *9*(3), 217–227. https://doi.org/10.1016/j.cmet.2009.01.010

Gnana-Prakasam, J. P., Martin, P. M., Smith, S. B., & Ganapathy, V. (2010). Expression and function of iron-regulatory proteins in retina. *IUBMB life*, *62*(5), 363–370. https://doi.org/10.1002/iub.326

Gudmundsson, S., Singer-Berk, M., Watts, N. A., Phu, W., Goodrich, J. K., Solomonson, M., Genome Aggregation Database Consortium, Rehm, H. L., MacArthur, D. G., & O'Donnell-Luria, A. (2022). Variant interpretation using population databases: Lessons from gnomAd. *Human Mutation*, *43*(8),1012–1030.19 https://doi.org/10.1002/humu.24309

Hesseling, P. B., Wessels, G., & van Riet, F. A. (1995). The Tygerberg Hospital Children's Tumour Registry 1983-1993. *European journal of cancer (Oxford, England : 1990)*, *31A*(9), 1471–1475. https://doi.org/10.1016/0959-8049(95)00289-u

Hunt, S. E., Moore, B., Amode, R. M., Armean, I. M., Lemos, D., Mushtaq, A., Parton, A., Schuilenburg, H., Szpak, M., Thormann, A., Perry, E., Trevanion, S. J., Flicek, P., Yates, A. D., & Cunningham, F. (2022). Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor-A tutorial. *Human Mutation*, *43*(8), 986–997. https://doi.org/10.1002/humu.24298

Indovina, P., Acquaviva, A., De Falco, G., Rizzo, V., Onnis, A., Luzzi, A., Giorgi, F., Hadjistilianou, T., Toti, P., Tomei, V., Pentimalli, F., Carugi, A., & Giordano, A. (2009).Downregulation and aberrant promoter methylation of *p16INK4A*: A possible novel heritable susceptibility marker to retinoblastoma. *Journal of Cellular Physiology*, 143-150. https://doi.org/10.1002/jcp.22019

Kato, M. V., Shimizu, T., Ishizaki, K., Kaneko, A., Yandell, D. W., Toguchida, J., & Sasaki, M. S. (1996). Loss of heterozygosity on chromosome 17 and mutation of the p53 gene in retinoblastoma. *Cancer Letters*, *106*(1), 75–82. https://doi.org/10.1016/0304-3835(96)04305-4

Ketteler, P., Hülsenbeck, I., Frank, M., Schmidt, B., Jöckel, K. H., & Lohmann, D. R. (2020). The impact of RB1 genotype on incidence of second tumours in heritable retinoblastoma. *European Journal of Cancer (Oxford, England : 1990)*, *133*, 47–55. https://doi.org/10.1016/j.ejca.2020.04.005

Knudson A. G., Jr (1971). Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, *68*(4), 820–823. https://doi.org/10.1073/pnas.68.4.820

Kooi, I. E., Mol, B. M., Massink, M. P., de Jong, M. C., de Graaf, P., van der Valk, P., Meijers-Heijboer, H., Kaspers, G. J., Moll, A. C., Te Riele, H., Cloos, J., & Dorsman, J. C. (2016). A Meta-Analysis of

Retinoblastoma Copy Numbers Refines the List of Possible Driver Genes Involved in Tumor Progression. *PloS one*, *11*(4), e0153323. https://doi.org/10.1371/journal.pone.0153323

Kubo, S., Yamamoto, H., Kajimura, N., Omori, Y., Maeda, Y., Chaya, T., & Furukawa, T. (2021). Functional analysis of Samd11, a retinal photoreceptor PRC1 component, in establishing rod photoreceptor identity. *Scientific Reports*, *11*(1), 4180. https://doi.org/10.1038/s41598-021-83781-1

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., Holmes, J. B., … Maglott, D. R. (2018). ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, *46*(D1), D1062–D1067. https://doi.org/10.1093/nar/gkx1153

Lee, W. H., Bookstein, R., Hong, F., Young, L. J., Shew, J. Y., & Lee, E. Y. (1987). Human retinoblastoma susceptibility gene: cloning, identification, and sequence. *Science (New York, N.Y.)*, *235*(4794), 1394–1399. https://doi.org/10.1126/science.3823889

Lee, W. H., Shew, J. Y., Hong, F. D., Sery, T. W., Donoso, L. A., Young, L. J., Bookstein, R., & Lee, E. Y. (1987). The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity. *Nature*, *329*(6140), 642–645. https://doi.org/10.1038/329642a0

Li, L., Li, B., Zhang, H., Bai, S., Wang, Y., Zhao, B., & Jonas, J. B. (2011). Lentiviral vector-mediated PAX6 overexpression promotes growth and inhibits apoptosis of human retinoblastoma cells. *Investigative Ophthalmology & Visual Science*, *52*(11), 8393–8400. https://doi.org/10.1167/iovs.11-8139

Lin, F. Y., & Chintagumpala, M. M. (2021). Neonatal Retinoblastoma. *Clinics in Perinatology*, *48*(1), 53–70. https://doi.org/10.1016/j.clp.2020.12.001

Livide, G., Epistolato, M. C., Amenduni, M., Disciglio, V., Marozza, A., Mencarelli, M. A., Toti, P., Lazzi, S., Hadjistilianou, T., De Francesco, S., D'Ambrosio, A., Renieri, A., & Ariani, F. (2012). Epigenetic and copy number variation analysis in retinoblastoma by MS-MLPA. *Pathology oncology research : POR*, *18*(3), 703–712. https://doi.org/10.1007/s12253-012-9498-8

Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., Bhurji, S. K., Bignell, A., Boddu, S., Branco Lins, P. R., Brooks, L., Ramaraju, S. B.,

Charkhchi, M., Cockburn, A., Da Rin Fiorretto, L., Davidson, C., … Flicek, P. (2023). Ensembl 2023. *Nucleic acids research*, *51*(D1), D933–D941. https://doi.org/10.1093/nar/gkac958

Maurya, R., Kumar, B., & Sundar, S. (2013). Evaluation of salt-out method for the isolation of DNA from whole blood: a pathological approach of DNA based diagnosis. *Int J Life Sci Biotechnol Pharma Res*, *2*(2), 53-57.

McEvoy, J. D., & Dyer, M. A. (2015). Genetic and Epigenetic Discoveries in Human Retinoblastoma. *Critical Reviews in Oncogenesis*, *20*(3-4), 217–225. https://doi.org/10.1615/critrevoncog.2015013711

McEvoy, J., Nagahawatte, P., Finkelstein, D., Richards-Yutz, J., Valentine, M., Ma, J., Mullighan, C., Song, G., Chen, X., Wilson, M., Brennan, R., Pounds, S., Becksfort, J., Huether, R., Lu, C., Fulton, R. S., Fulton, L. L., Hong, X., Dooling, D. J., Ochoa, K., … Dyer, M. A. (2014). RB1 gene inactivation by chromothripsis in human retinoblastoma. *Oncotarget*, *5*(2), 438–450. https://doi.org/10.18632/oncotarget.1686

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122. https://doi.org/10.1186/s13059-016-0974-4

Meng, B., Wang, Y., & Li, B. (2014). Suppression of PAX6 promotes cell proliferation and inhibits apoptosis in human retinoblastoma cells. *International Journal of Molecular Medicine*, *34*(2), 399–408. https://doi.org/10.3892/ijmm.2014.1812

Ndlovu, S., Esterhuizen, T., Uys, R., Van Zyl, A., & Kruger, M. (2023). Trends in childhood cancers at Tygerberg Hospital from 1994 to 2014. *South African Journal of Child Health*, 96–98. https://doi.org/10.7196/sajch.2023.v17i2.1944

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research*, *31*(13), 3812–3814. https://doi.org/10.1093/nar/gkg509

Nork, T. M., Poulsen, G. L., Millecchia, L. L., Jantz, R. G., & Nickells, R. W. (1997). p53 regulates apoptosis in human retinoblastoma. *Archives of Ophthalmology (Chicago, Ill. : 1960)*, *115*(2), 213–219. https://doi.org/10.1001/archopht.1997.01100150215011

Plousiou, M., & Vannini, I. (2019). Non-Coding RNAs in Retinoblastoma. *Frontiers in Genetics*, *10*, 1155. https://doi.org/10.3389/fgene.2019.01155

Reis, A. H., Vargas, F. R., & Lemos, B. (2012). More epigenetic hits than meets the eye: microRNAs and genes associated with the tumorigenesis of retinoblastoma. *Frontiers in Genetics*, *3*, 284. https://doi.org/10.3389/fgene.2012.00284

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

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., & 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), 405–424. https://doi.org/10.1038/gim.2015.30

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Roelfsema, J. H., & Peters, D. J. (2007). Rubinstein-Taybi syndrome: clinical and molecular overview. *Expert Reviews in Molecular Medicine*, *9*(23), 1–16. https://doi.org/10.1017/S1462399407000415

Rushlow, D., Piovesan, B., Zhang, K., Prigoda-Lee, N. L., Marchong, M. N., Clark, R. D., & Gallie, B. L. (2009). Detection of mosaic RB1 mutations in families with retinoblastoma. *Human Mutation*, *30*(5), 842–851. https://doi.org/10.1002/humu.20940

Rushlow, D. E., Mol, B. M., Kennett, J. Y., Yee, S., Pajovic, S., Thériault, B. L., Prigoda-Lee, N. L., Spencer, C., Dimaras, H., Corson, T. W., Pang, R., Massey, C., Godbout, R., Jiang, Z., Zacksenhaus, E., Paton, K., Moll, A. C., Houdayer, C., Raizis, A., Halliday, W., … Gallie, B. L. (2013). Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *The Lancet. Oncology*, *14*(4), 327–334. https://doi.org/10.1016/S1470-2045(13)70045-7

Shinohara, E. T., DeWees, T., & Perkins, S. M. (2014). Subsequent malignancies and their effect on survival in patients with retinoblastoma. *Pediatric Blood & Cancer*, *61*(1), 116–119. https://doi.org/10.1002/pbc.24714

Singh, U., Malik, M. A., Goswami, S., Shukla, S., & Kaur, J. (2016). Epigenetic regulation of human retinoblastoma. *Tumour biology : the Journal of the International Society for Oncodevelopmental Biology and Medicine*, *37*(11), 14427–14441. https://doi.org/10.1007/s13277-016-5308-3

Skowron, P., Farooq, H., Cavalli, F. M. G., Morrissy, A. S., Ly, M., Hendrikse, L. D., Wang, E. Y., Djambazian, H., Zhu, H., Mungall, K. L., Trinh, Q. M., Zheng, T., Dai, S., Stucklin, A. S. G., Vladoiu, M. C., Fong, V., Holgado, B. L., Nor, C., Wu, X., Abd-Rabbo, D., … Taylor, M. D. (2021). The transcriptional landscape of Shh medulloblastoma. *Nature Communications*, *12*(1), 1749. https://doi.org/10.1038/s41467-021-21883-0

Smith, C. G., Naven, M., Harris, R., Colley, J., West, H., Li, N., Liu, Y., Adams, R., Maughan, T. S., Nichols, L., Kaplan, R., Wagner, M. J., McLeod, H. L., & Cheadle, J. P. (2013). Exome resequencing identifies potential tumor-suppressor genes that predispose to colorectal cancer. *Human Mutation*, *34*(7), 1026–1034. https://doi.org/10.1002/humu.22333

Sparks, A. B., Morin, P. J., Vogelstein, B., & Kinzler, K. W. (1998). Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer research*, *58*(6), 1130–1134.

Steinhaus, R., Proft, S., Schuelke, M., Cooper, D. N., Schwarz, J. M., & Seelow, D. (2021). MutationTaster2021. *Nucleic Acids Research*, *49*(W1), W446–W451. https://doi.org/10.1093/nar/gkab266

Stenfelt, S., Blixt, M. K. E., All-Ericsson, C., Hallböök, F., & Boije, H. (2017). Heterogeneity in retinoblastoma: a tale of molecules and models. *Clinical and Translational Medicine*, *6*(1), 42. https://doi.org/10.1186/s40169-017-0173-2

Stuart, K. V., Shepherd, D. J., Kruger, M., & Singh, E. (2022). The Incidence of Retinoblastoma in South Africa: Findings from the South African National Cancer Registry (2004-2018). *Ophthalmic Epidemiology*, *29*(6), 681–687. https://doi.org/10.1080/09286586.2021.2013900

1000 Genomes Project Consortium, Auton, A., Brooks, L. D., Durbin, R. M., Garrison, E. P., Kang, H. M., Korbel, J. O., Marchini, J. L., McCarthy, S., McVean, G. A., & Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

Ueki, Y., & Reh, T. A. (2012). Activation of BMP-Smad1/5/8 signaling promotes survival of retinal ganglion cells after damage in vivo. *PloS one*, *7*(6), e38690. https://doi.org/10.1371/journal.pone.0038690

Vasu, U., Nithyanandam, S., D'Souza, S., & Kamath, S. (2011). Retinoblastoma in patients with regressed retinopathy of prematurity. *Indian journal of ophthalmology*, *59*(6), 501–502. https://doi.org/10.4103/0301-4738.86322

Xu, X. L., Fang, Y., Lee, T. C., Forrest, D., Gregory-Evans, C., Almeida, D., Liu, A., Jhanwar, S. C., Abramson, D. H., & Cobrinik, D. (2009). Retinoblastoma has properties of a cone precursor tumor and depends upon cone-specific MDM2 signaling. *Cell*, *137*(6), 1018–1031. https://doi.org/10.1016/j.cell.2009.03.051

Xu, L., Shen, L., Polski, A., Prabakar, R. K., Shah, R., Jubran, R., Kim, J. W., Biegel, J., Kuhn, P., Cobrinik, D., Hicks, J., Gai, X., & Berry, J. L. (2020). Simultaneous identification of clinically relevant *RB1* mutations and copy number alterations in aqueous humor of retinoblastoma eyes. *Ophthalmic Genetics*, *41*(6), 526–532. https://doi.org/10.1080/13816810.2020.1799417

Zhang, J., Vlasevska, S., Wells, V. A., Nataraj, S., Holmes, A. B., Duval, R., Meyer, S. N., Mo, T., Basso, K., Brindle, P. K., Hussein, S., Dalla-Favera, R., & Pasqualucci, L. (2017). The CREBBP Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discovery*, *7*(3), 322–337. https://doi.org/10.1158/2159-8290.CD-16-1417

Appendices

Appendix A:

Human Research Ethics Committee (Medical), University of the Witwatersrand Ethics Certificate (M180855)

Appendix B:

Human Research Ethics Committee (Medical), University of the Witwatersrand Ethics Certificate (M230782)

Appendix C:

RB1 sequence variants in all patients

Appendix D

Extended list of variants of interest

Appendix E

BAM images to evaluate variant coverage

Appendix F:

Information, Consent and Assent forms

Appendix G:

Title Approval and Approved Research Protocol

Appendix H:

Plagiarism declaration and Turnitin report

Appendix I:

Journal Author Guidelines for Submissible format

R14/49 Dr L Lamola, et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180855

NAME: Dr L Lamola, et al
(Principal Investigator)
DEPARTMENT: School of Pathology
Department of Human Genetics
National Health Laboratory Service

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

DATE CONSIDERED: 31/08/2018

DECISION: Approved unconditionally

CONDITIONS:

SUPERVISOR: Not applicable

APPROVED BY: CB Penny
Dr CB Penny, Chairperson, HREC (Medical)

DATE OF APPROVAL: 26/02/2019

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

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

I/We fully understand the conditions under which I am/we are authorised 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 resubmit to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date of the meeting when the study was initially reviewed. In this case, the study was initially reviewed in **August** and will therefore reports and re-certification will be due early in the month of **August** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

26/02/2019
Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

R49 Ms D Beukman

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M230782

NAME: (Principal Investigator)	Ms D Beukman
DEPARTMENT:	School of Pathology Division of Human Genetics Medical School University
PROJECT TITLE:	*Beyond RB1: Determining genetic causes of retinablastoma in South African patients*
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M18/08/55
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Dr L Lamola; Ms N Mayevu
APPROVED BY:	Professor P Ruff, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2023/08/18 **EXPIRY DATE:** 2028/08/17

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**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **July** and therefore reports and re-certification will be due in the month of **July** each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator _____ Date _____

Appendix C:

RB1 sequence variants in all patients

Appendix 3 Table 5: *RB1* Exonic variants

Location	Consequence	HGVSc	HGVSp	Existing variant	SIFT	Polyphen-2	CADD_ PHRED	AF [a]	CLIN_SIG [b]
PC26									
13:48955458-48955458	missense variant	NM_000321.3:c.1574C>G	NP_000312.2:p.Ala525Gly	rs4151539	tolerated(0.07)	probably damaging(0.962)	24.7	0.0062	benign/likely benign,
PC43									
13:48878084-48878093	Inframe deletion	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	rs572454921			13.38	0.0108	benign
13:48955516-48955516	synonymous	NM_000321.3:c.1632A>G	NP_000312.2:p.Arg544%3D	rs143948310			12.48	0.0020	Benign, likely benign
PC50									
13:48878084-48878093	Inframe deletion	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	rs572454921			13.38	0.0108	benign
PC79 No exonic variants									
PC90 No exonic variants									
PC104 No exonic variants									
PC120 No exonic variants									
PC123									
13:48878084-48878093	Inframe deletion	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	rs572454921			13.38	0.0108	benign

All exonic sequence variants in *RB1* with a minor allele frequency lower than 5 %.

[a] Allele frequency as reported in gnomAD

[b] Clinical significance as reported in Clinvar and ACMG/AMP variant classification codes

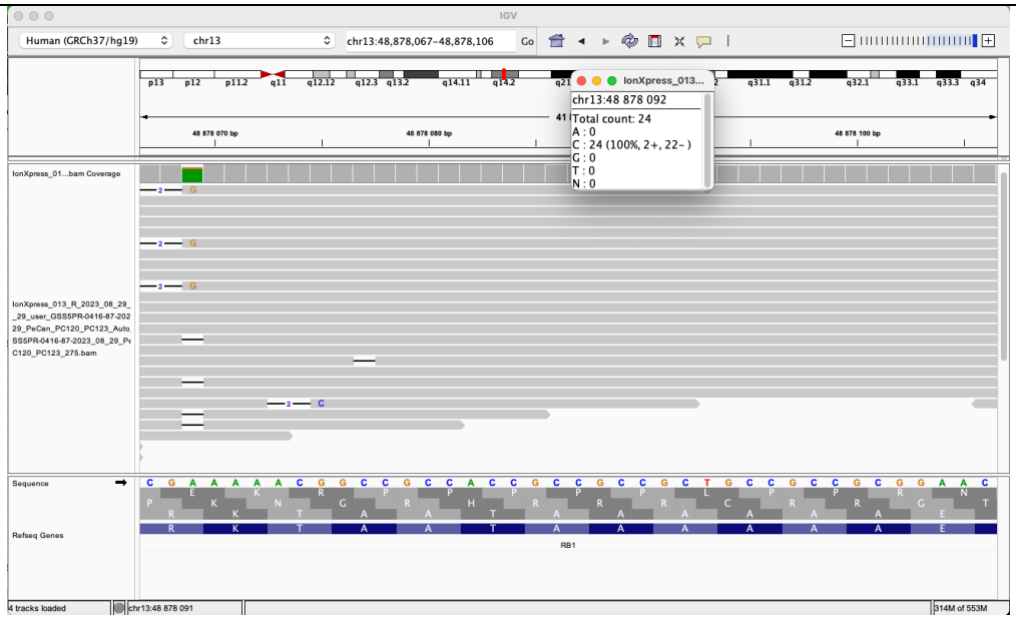
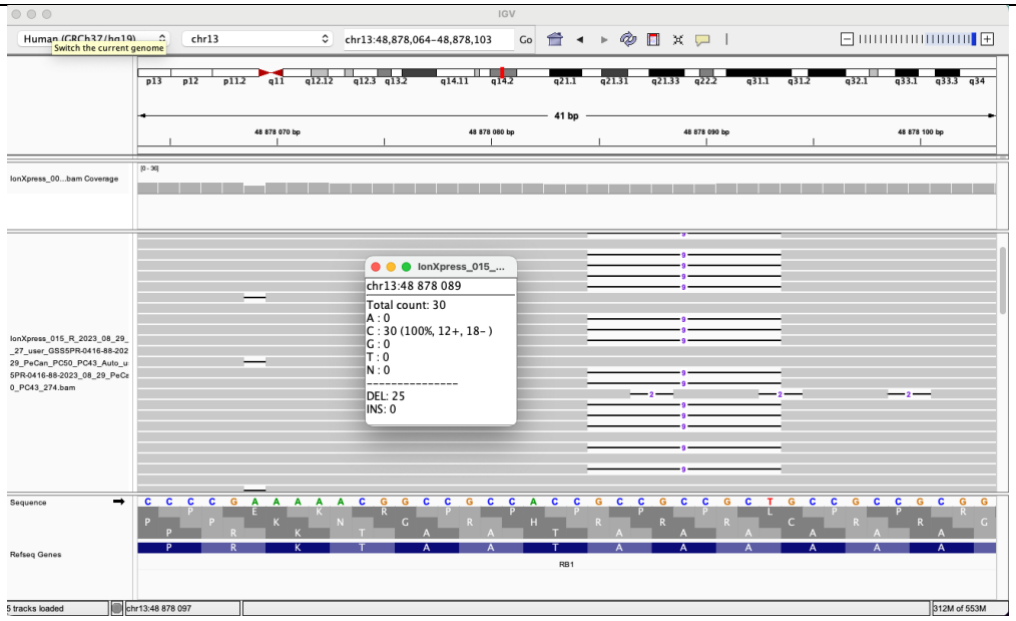
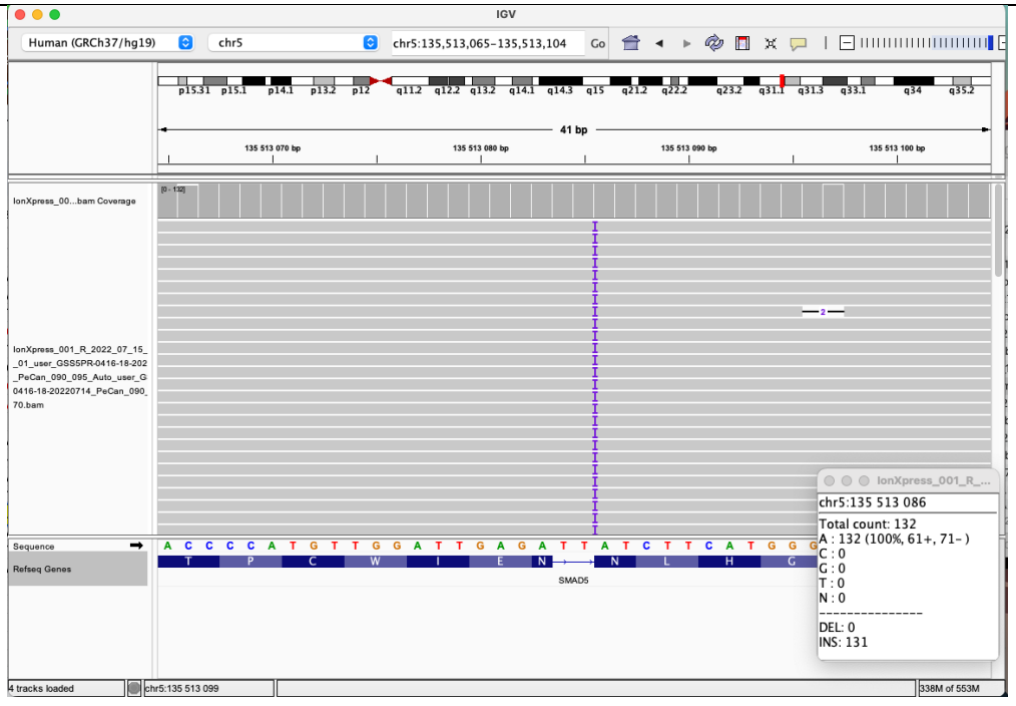
Appendix D: Extended list of variants of interest

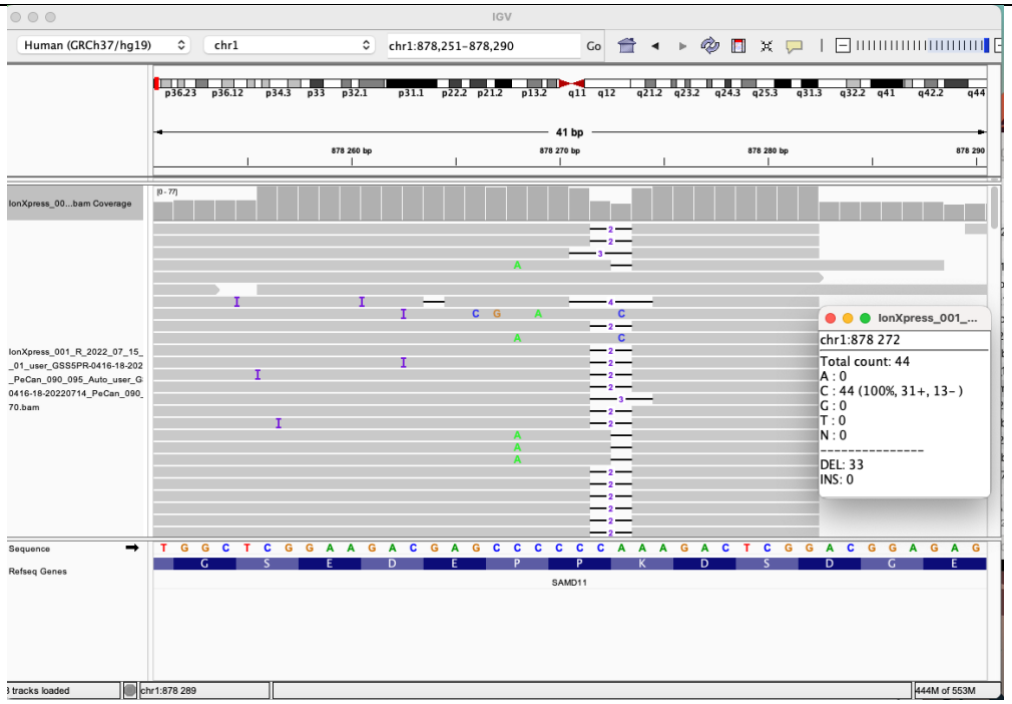
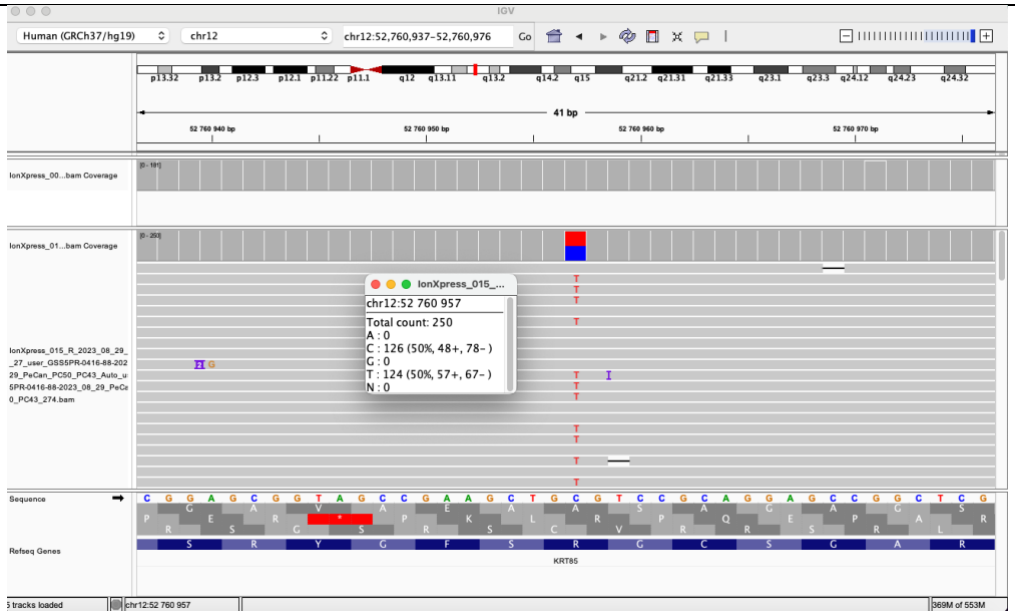
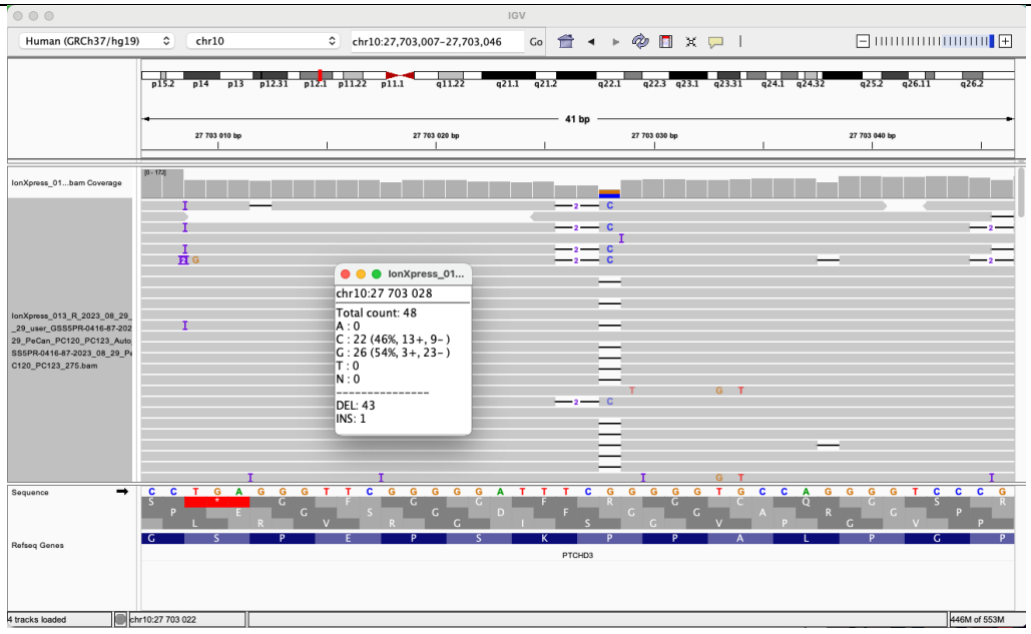
	A	B	C	D	E	F	G	H	I	J
1	**Appendix D**									
2	Location	Allele	Consequence	IMPACT	SYMBOL	Gene	HGVSc	HGVSp	cDNA_position	Existing_variation
3	**PC26**									
4	1:879381-879381	T	missense_variant	MODERATE	SAMD11	148398	NM_001385640.1:c.2386C>T	NP_001372569.1:p.Leu796Phe	2895	rs116362966
5	13:48955458-48955458	G	missense_variant	MODERATE	RB1	5925	NM_000321.3:c.1574C>G	NP_000312.2:p.Ala525Gly	1736	rs4151539
6	5:112173899-112173899	T	missense_variant	MODERATE	APC	324	NM_000038.6:c.2608C>T	NP_000029.2:p.Pro870Ser	2667	rs33974176
7	9:98209623-98209625	GT	frameshift_variant	HIGH	PTCH1	5727	NM_000264.5:c.3913del	NP_000255.2:p.Asp1305ThrfsTer67	4818-4820	-
8	12:52760957-52760957	T	missense_variant	MODERATE	KRT85	3891	NM_002283.4:c.233G>A	NP_002274.1:p.Arg78His	309	rs61630004
9	19:5831602-5831602	T	missense_variant	MODERATE	FUT6	2528	NM_000150.3:c.977G>A	NP_000141.1:p.Arg326Gln	1233	rs61739552
10	19:5832077-5832077	G	frameshift_variant	HIGH	FUT6	2528	NM_000150.3:c.501_502insC	NP_000141.1:p.Tyr168LeufsTer230	757-758	rs150560931
11	14:21811196-21811196	G	missense_variant,splic	MODERATE	RPGRIP1	57096	NM_001377523.1:c.1319A>G	NP_001364452.1:p.Asp440Gly	1385	rs17103671
12	3:37083760-37083760	A	missense_variant,splic	MODERATE	MLH1	4292	NM_001167617.3:c.1375G>A	NP_001161089.1:p.Glu459Lys	1921	rs63751244
13	2:234668955-234668955	A	missense_variant	MODERATE	UGT1A1	54658	NM_000463.3:c.22G>A	NP_000454.1:p.Gly8Arg	40	rs749552053
14	5:79950727-79950736	-	inframe_deletion	MODERATE	MSH3	4437	NM_002439.5:c.195_203del	NP_002430.3:p.Pro67_Pro69del	258-266	rs60484572
15	22:29130456-29130456	A	missense_variant	MODERATE	CHEK2	11200	NM_001005735.2:c.254C>T	NP_001005735.1:p.Pro85Leu	312	rs17883862
16	**PC43**									
17	1:878271-878273	-	frameshift_variant	HIGH	SAMD11	148398	NM_001385640.1:c.1890_1891	NP_001372569.1:p.Lys631ArgfsTer71	2399-2400	-
18	1:879481-879481	C	missense_variant	MODERATE	SAMD11	148398	NM_001385640.1:c.2486G>C	NP_001372569.1:p.Gly829Ala	2995	rs113383096
19	13:48878084-48878093	-	inframe_deletion	MODERATE	RB1	5925	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	199-207	rs572454921
20	16:3778363-3778363	T	missense_variant	MODERATE	CREBBP	1387	NM_004380.3:c.6685G>A	NP_004371.2:p.Gly2229Ser	7482	rs139688311
21	3:39307063-39307063	A	missense_variant	MODERATE	CX3CR1	1524	NM_001171172.2:c.938C>T	NP_001164643.1:p.Ala313Val	1140	rs137947370
22	4:88533730-88533730	C	missense_variant	MODERATE	DSPP	1834	NM_014208.3:c.392T>C	NP_055023.2:p.Ile131Thr	512	rs61731009
23	12:52760957-52760957	T	missense_variant	MODERATE	KRT85	3891	NM_002283.4:c.233G>A	NP_002274.1:p.Arg78His	309	rs61630004
24	10:54531361-54531362	CC	missense_variant	MODERATE	MBL2	4153	NM_000242.3:c.34_35delinsGG	NP_000233.1:p.Leu12Gly	51-52	-
25	14:21816432-21816432	A	missense_variant	MODERATE	RPGRIP1	57096	NM_001377523.1:c.1697G>A	NP_001364452.1:p.Gly566Glu	1763	rs34725281
26	16:2133726-2133726	T	missense_variant	MODERATE	TSC2	7249	NM_000548.5:c.3914C>T	NP_000539.2:p.Pro1305Leu	4024	rs45517320
27	13:32953529-32953529	T	missense_variant	MODERATE	BRCA2	675	NM_000059.4:c.8830A>T	NP_000050.3:p.Ile2944Phe	9029	rs4987047
28	3:37092025-37092025	T	missense_variant	MODERATE	MLH1	4292	NM_001258271.2:c.1945C>T	NP_001245200.1:p.His649Tyr	1975	rs2020873
29	2:234545330-234545330	T	missense_variant	MODERATE	UGT1A10	54575	NM_019075.2:c.162G>T	NP_061948.1:p.Glu54Asp	208	rs148859354
30	16:2097712-2097712	A	missense_variant,splic	MODERATE	NTHL1	4913	NM_002528.7:c.113C>T	NP_002519.2:p.Ala38Val	124	rs202082304
31	**PC50**									
32	1:879481-879481	C	missense_variant	MODERATE	SAMD11	148398	NM_001385640.1:c.2486G>C	NP_001372569.1:p.Gly829Ala	2995	rs113383096
33	13:48878084-48878093	-	inframe_deletion	MODERATE	RB1	5925	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	199-207	rs572454921
34	5:112154980-112154980	G	missense_variant	MODERATE	APC	324	NM_000038.6:c.1251T>G	NP_000029.2:p.Cys417Trp	1310	-
35	5:135513085-135513085	C	frameshift_variant,spli	HIGH	SMAD5	4090	NM_001001419.3:c.1313-1_131	NP_001001419.1:p.Asn438ThrfsTer11	1758-1759	-
36	12:52760957-52760957	T	missense_variant	MODERATE	KRT85	3891	NM_002283.4:c.233G>A	NP_002274.1:p.Arg78His	309	rs61630004
37	17:41243800-41243800	T	missense_variant	MODERATE	BRCA1	672	NM_007294.4:c.3748G>A	NP_009225.1:p.Glu1250Lys	3861	rs28897686
38	16:3779210-3779217	GGGGGG	frameshift_variant	HIGH	CREBBP	1387	NM_001079846.1:c.5723del	NP_001073315.1:p.Pro1908HisfsTer30	5921-5927	rs587783507
39	**PC79**									
40	4:88534326-88534326	A	missense_variant	MODERATE	DSPP	1834	NM_014208.3:c.988G>A	NP_055023.2:p.Ala330Thr	1108	rs201942511
41	9:98231233-98231233	T	missense_variant	MODERATE	PTCH1	5727	NM_001083606.3:c.1597G>A	NP_001077075.1:p.Glu533Lys	2048	rs62637629
42	5:135513085-135513085	C	frameshift_variant,spli	HIGH	SMAD5	4090	NM_001001419.3:c.1313-1_131	NP_001001419.1:p.Asn438ThrfsTer11	1758-1759	-
43	12:52758810-52758810	T	missense_variant	MODERATE	KRT85	3891	NM_002283.4:c.565G>A	NP_002274.1:p.Asp189Asn	641	rs112554450
44	6:49587034-49587034	G	missense_variant	MODERATE	RHAG	6005	NM_000324.3:c.199T>C	NP_000315.2:p.Phe67Leu	226	rs116356543
45	2:47637246-47637246	G	missense_variant	MODERATE	MSH2	4436	NM_000251.3:c.380A>G	NP_000242.1:p.Asn127Ser	416	rs17217772

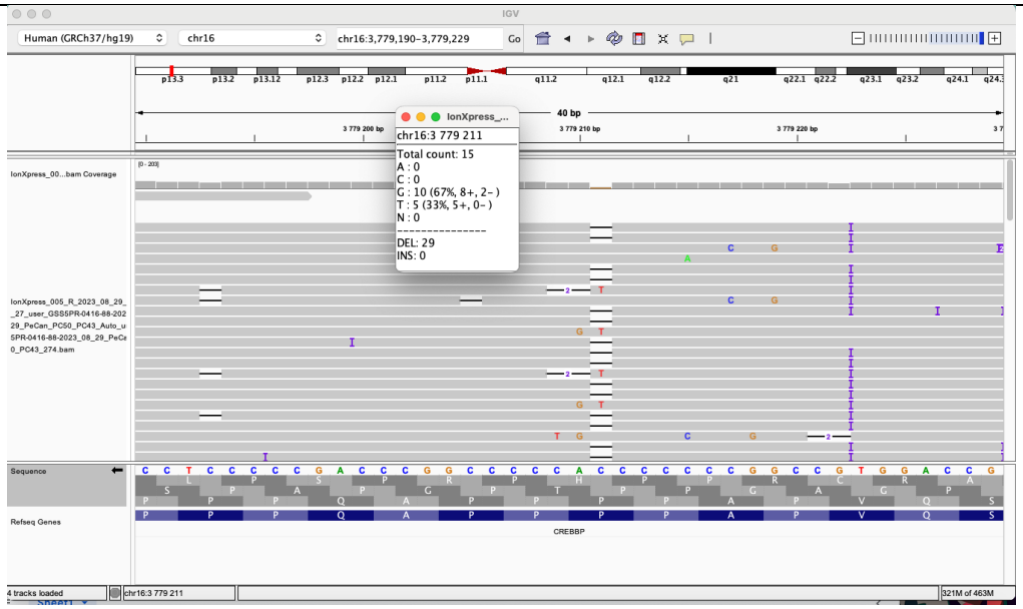
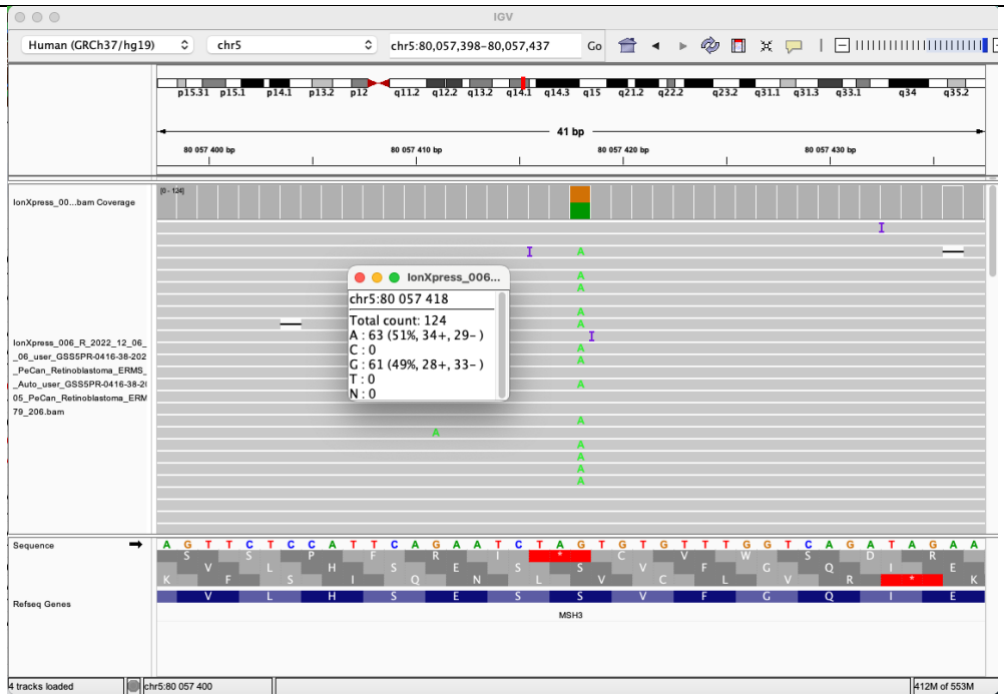
	A	B	C	D	E	F	G	H	I	J
46	5:80057418-80057418	A	missense_variant	MODERATE	MSH3	4437	NM_002439.5:c.1817G>A	NP_002430.3:p.Ser606Asn	1893	rs34168832
47	**PC90**									
48	1:878271-878273	-	frameshift_variant	HIGH	SAMD11	148398	NM_001385640.1:c.1890_1891	NP_001372569.1:p.Lys631ArgfsTer71	2399-2400	-
49	7:138430019-138430019	A	missense_variant	MODERATE	ATP6V0A4	50617	NM_020632.3:c.1327A>T	NP_065683.2:p.Asn443Tyr	1610	rs542662856
50	4:88534140-88534140	T	missense_variant	MODERATE	DSPP	1834	NM_014208.3:c.802G>T	NP_055023.2:p.Gly268Trp	922	rs61738508
51	5:135513085-135513085	C	frameshift_variant,spli	HIGH	SMAD5	4090	NM_001001420.3:c.1313-1_131	NP_001001420.1:p.Asn438ThrfsTer11	1599-1600	-
52	19:5832203-5832203	A	missense_variant	MODERATE	FUT6	2528	NM_000150.3:c.376C>T	NP_000141.1:p.Arg126Trp	632	rs111589452
53	12:52758810-52758810	T	missense_variant	MODERATE	KRT85	3891	NM_002283.4:c.565G>A	NP_002274.1:p.Asp189Asn	641	rs112554450
54	17:41243502-41243502	A	missense_variant	MODERATE	BRCA1	672	NM_007300.4:c.4046C>T	NP_009231.2:p.Thr1349Met	4159	rs80357345
55	**PC104**									
56	1:878271-878273	-	frameshift_variant	HIGH	SAMD11	148398	NM_001385640.1:c.1890_1891	NP_001372569.1:p.Lys631ArgfsTer71	2399-2400	-
57	5:79950707-79950725	-	inframe_deletion	MODERATE	MSH3	4437	NM_002439.5:c.162_179del	NP_002430.3:p.Ala57_Ala62del	238-255	-
58	**PC120**									
59	5:112173899-112173899	T	missense_variant	MODERATE	APC	324	NM_000038.6:c.2608C>T	NP_000029.2:p.Pro870Ser	2667	rs33974176
60	4:88533730-88533730	C	missense_variant	MODERATE	DSPP	1834	NM_014208.3:c.392T>C	NP_055023.2:p.Ile131Thr	512	rs61731009
61	6:47846362-47846362	C	missense_variant	MODERATE	PTCHD4	442213	NM_001013732.4:c.2218A>G	NP_001013754.3:p.Thr740Ala	2393	rs113836634
62	10:27703027-27703028	-	frameshift_variant	HIGH	PTCHD3	374308	NM_001034842.4:c.152del	NP_001030014.2:p.Pro51ArgfsTer69	249	-
63	5:135513085-135513085	C	frameshift_variant,spli	HIGH	SMAD5	4090	NM_001001419.3:c.1313-1_131	NP_001001419.1:p.Asn438ThrfsTer11	1758-1759	-
64	12:69218490-69218490	T	intron_variant	MODIFIER	MDM2	4193	NM_001145337.3:c.505+59G>T	-	-	rs148010327
65	13:32953529-32953529	T	missense_variant	MODERATE	BRCA2	675	NM_000059.4:c.8830A>T	NP_000050.3:p.Ile2944Phe	9029	rs4987047
66	2:234545898-234545898	A	missense_variant	MODERATE	UGT1A10	54575	NM_019075.2:c.730C>A	NP_061948.1:p.Leu244Ile	776	rs28969685
67	16:2096377-2096377	C	splice_polypyrimidine_	LOW	NTHL1	4913	NM_001318193.2:c.116-10C>G	-	-	rs3211970
68	**PC123**									
69	1:865584-865584	A	missense_variant	MODERATE	SAMD11	148398	NM_001385640.1:c.659G>A	NP_001372569.1:p.Arg220Gln	1168	rs148711625
70	1:878271-878273	-	frameshift_variant	HIGH	SAMD11	148398	NM_001385640.1:c.1890_1891	NP_001372569.1:p.Lys631ArgfsTer71	2399-2400	-
71	1:879381-879381	T	missense_variant	MODERATE	SAMD11	148398	NM_001385640.1:c.2386C>T	NP_001372569.1:p.Leu796Phe	2895	rs116362966
72	13:48878084-48878093	-	inframe_deletion	MODERATE	RB1	5925	NM_000321.3:c.45_53del	NP_000312.2:p.Ala16_Ala18del	199-207	rs572454921
73	6:36651889-36651889	T	missense_variant	MODERATE	CDKN1A	1026	NM_000389.5:c.11C>T	NP_000380.1:p.Pro4Leu	101	rs4986866
74	4:88533730-88533730	C	missense_variant	MODERATE	DSPP	1834	NM_014208.3:c.392T>C	NP_055023.2:p.Ile131Thr	512	rs61731009
75	5:176524540-176524540	T	missense_variant	MODERATE	FGFR4	2264	NM_001291980.2:c.2068C>T	NP_001278909.1:p.Arg690Cys	2240	rs148143006
76	6:26091149-26091149	A	missense_variant	MODERATE	HFE	3077	NM_139006.3:c.157G>A	NP_620575.1:p.Val53Met	169	rs28934889
77	6:26091149-26091149	A	missense_variant	MODERATE	HFE	3077	NM_139009.3:c.88G>A	NP_620578.1:p.Val30Met	100	rs28934889
78	9:37006494-37006494	T	missense_variant	MODERATE	PAX5	5079	NM_001280552.2:c.451G>A	NP_001267481.1:p.Val151Ile	688	rs115889954
79	14:94847262-94847262	A	missense_variant	MODERATE	SERPINA1	5265	NM_001127700.2:c.863A>T	NP_001121172.1:p.Glu288Val	1083	rs17580
80	16:2134418-2134418	A	missense_variant	MODERATE	TSC2	7249	NM_001114382.3:c.4126G>A	NP_001107854.1:p.Gly1376Arg	4236	rs45466399

Appendix E

BAM images to evaluate variant coverage

RB1 Missense variant (Likely benign) NM_000321.3:c.1574C>G NP_000312.2:p.Ala525Gly	
RB1 Inframe deletion (Benign) NM_000321.3:c.45_53del NP_000312.2:p.Ala16_Ala18del	
SMAD5 insertion Frame shift variant (VUS) NM_001001419.3:c.1313-1_1313insC NP_001001419.1:p.Asn438ThrfsTer11	

SAMD11 deletion Frame shift variant (VUS) NM_001385640.1:c.1890_1891del NP_001372569.1:p.Lys631ArgfsTer71	
KRT85 Missense variant (Likely benign) NM_002283.4:c.233G>A NP_002274.1:p.Arg78His	
PTCHD3 deletion Frame shift variant (VUS) NM_001034842.4:c.152del NP_001030014.2:p.Pro51ArgfsTer69	

CREBBP deletion Frameshift variant (VUS) NM_001079846.1:c.5723del NP_001073315.1:p.Pro1908HisfsTer30	
MSH3 Missense variant (VUS) NM_002439.5:c.1817G>A NP_002430.3:p.Ser606Asn	

Appendix F:

Information, Consent and Assent forms

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 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

Protocol: M180855

Inherited cancer predisposing mutations in a cohort of childhood cancer patients

CONSENT FORM – PARENTAL CONSENT FOR CHILD PARTICIPANTS

PARTICIPANT STUDY ID:

The purpose of this study, procedures to be followed, risks and benefits, have been explained to me and my child. I have been allowed to ask questions, and my questions have been answered to my satisfaction. I have been told whom to contact if I have questions, to discuss problems, concerns, or suggestions related to the research, or to obtain information or offer input about the research. I have read the attached information sheet this and agree to allow my child to participate in this research study with the understanding that I may withdraw my child at any time. To the best of my ability, I have discussed the study with my child, who agrees to be in the study. I have explained to my child that he/she may approach the researchers to be re-consented when he/she turns 18 if he/she wishes and that he/she may withdraw from the study at any time. I have been told that I will be given a signed and dated copy of this consent form if I would like one.

	Please initial
I understand that the purpose of this consent form is for me to inform the study what they can or cannot do with my child's samples, clinical information and photos.	
I understand that all procedures/tests on the samples have been or will be approved by the Human Research Ethics Committee of the University of the Witwatersrand. I understand that every time a new study is done on my child's sample, permission will be obtained from this Ethics Committee for the study to make sure that it is used only for the purposes stated above.	
I understand that if I have any questions at any time, they will be answered.	
I understand that confidentiality will be maintained at all times through the use of a code and numbering system.	
I understand that I may withdraw my child from the study at any time.	
I have discussed the study with my child.	

I am in agreement that my child's samples may be stored and used for the purposes described in the information sheet.	
I understand that I can choose to have my child's sample destroyed at any time.	
I understand that my child may not benefit directly from the research done on their sample.	
If this research does identify information that is relevant to my child's cancer, I would like to be contacted to receive the information.	Yes ☐ No ☐
If this research does identify genetic information that would be important or of health benefit to my family, or me, I would like to be contacted to receive the information.	Yes ☐ No ☐

Researcher	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Parent Participant	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)

Chairperson: Prof Eric Buch Acting CEO: Dr Karmani Chetty
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

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 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

Protocol: M180855

Dear Potential Participant,

Hello, we are Dr. Lindiwe Lamola, Dr. Candice Feben and Prof. Amanda Krause from the Division of Human Genetics at the University of the Witwatersrand and National Health Laboratory Service. We are doing research on childhood cancers. We work with other members of the Division of Human Genetics to study the genetic factors that cause childhood cancers in South Africans.

As we are trying to understand the genetic causes of cancer, we would like to invite your child to take part in this study, because your child has cancer. Participation in the study will mean answering some questions about the health and background of your child and giving a blood and/or saliva sample for genetic studies. We will also examine your affected child, review their clinical file and investigations, and take a blood and/ or saliva sample from them.

What are childhood cancers?

Childhood cancers are the cancers that develop in children from birth to 18 years of age. Childhood cancer is not just one disease, but a group of several major types of disease. There are many types of childhood cancers; these may include; neuroblastoma, retinoblastoma, Wilms tumor, and cancers of the brain, bone, and soft tissue. The childhood cancers are not common and differ from adult cancers in the way they grow, spread and how they are treated.

What is cancer?

Cancer is an abnormal growth of cells. Cells are small basic building blocks of all living things. All parts of the body are made up of cells. Cancer develops when damaged cells grow and spread faster than normal healthy cells.

What causes childhood cancers?

Childhood cancers are complex diseases and doctors sometimes aren't sure why some children get cancer and others don't. However, about 10% (1 in 10) of all cancers in children are caused by inherited changes in your genes – inherited means these changes can be passed from parents to children. For this study, we are interested in studying the genetic causes of childhood cancers in order to understand their role in the development of cancer.

What will the study involve?

In this study, we would like to understand the factors which may have played a part in your child developing cancer. To do this, we want to learn more about your family, specifically whether other family members have had cancer, and to learn more about your child, specifically related to his/her type of cancer, how old the child was when the cancer was diagnosed and whether there are any other problems which your child may have been born with or may have developed over time.

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

We would like to take a full family history from you and then perform a clinical examination of your child, looking at his/her weight and growth, skin, chest, abdomen, limbs and spine. The family history and examination will take approximately half an hour. We would like to look at the records in your child's hospital file to make sure that we collect all the relevant information

In order to study the genes, we will take a sample of blood and/or saliva from your child. Our blood and saliva contains cells. All the cells in our bodies contain a chemical called DNA that we inherit from our parents. DNA carries the instructions that your body must follow in order to exist and function. It contains all your genes. Your cells also contain a chemical called RNA. RNA is formed from DNA and helps DNA to do its job.

We can take the DNA and RNA out of the cells that are in the blood or the saliva sample that we take from your child. Once we have isolated the DNA and/or RNA we can study the genes. In this study, we may need to examine all of the genes to try and find the gene that may be the cause of your child's cancer.

What can you expect from the study?

Once you understand what the study is about and are happy to take part, we will ask you to sign a consent form on behalf of your child to join the study. If your child is old enough and able to understand the study, they may sign an assent form. A copy of these forms will be given to you to keep.

Once you have signed the consent form and your child has signed the assent form we will ask you some questions about your child's background and health. A complete clinical examination will be performed on your affected child. We may ask to access data relevant to your child's diagnosis from his/her clinical file or test results. This information will help us to understand your child's condition better and to interpret the data generated by our research.

A nursing sister or doctor will take up to three teaspoons (15ml) of blood from your child. Venupunctures (drawing blood) are normally done as part of routine medical care and present a slight risk of discomfort. Drawing blood may result in faintness, inflammation of the vein, pain, bruising or bleeding at the puncture site. Your protection is that experienced staff perform the procedures under sterile conditions.

We will also ask for a saliva (spit) sample from your affected child. A nursing sister or doctor will use a small, soft sponge to collect the saliva sample. The sponge is placed inside your child's mouth against the cheek and gently moved along the gums and cheek to soak up as much saliva as possible. He/she will then place the saliva in a special container and might repeat the process if we do not get enough saliva the first time. This procedure is painless and harmless.

Chairperson: Prof Eric Buch Acting CEO: Dr Karmani Chetty
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

Long-term storing of samples and data

Once we have done the research that we are planning for this project, we would like to store your child's DNA, genomic information and data on their health history and laboratory test results for future research on childhood cancer, and we will ask you to sign a separate consent form to give us permission to do this. The sample will be stored together with other participants' samples. The samples will be used in future childhood cancer studies.

Confidentiality

Your child's samples will be stored in a locked fridge and the personal information will be stored on a secure computer in a locked office. Only the investigators of this study will have access your child's name and personal details. Each participant will be given a code and number for the study so that no one knows who you are. The blood, saliva, RNA and DNA samples and the associated data will not be attached to your names – they will only have their code and number on them.

What are my rights as a participant?

You and your child do not have to join this study. Whether he/she decides to join this study or not will not affect your treatment in our clinic. You can also decide to leave the study at any time. You may choose to have your child's biological samples (blood, DNA, RNA and saliva) destroyed at any time. If you do not want your child to be part of this study or decide to leave the study, it will not affect your child's regular treatments or medical care in any way. We invite your child to approach the researchers to be re-consented when he/she turns 18 if he/she wishes. He/she can also decide to leave the study at any time.

You may not get any benefit directly from this study, but we hope it may help others who are affected by childhood cancers in the future. We may contact you in future if we believe we have identified the cause of your child's condition if you give your permission for this. Please understand that what we are trying to do is very difficult and could take a long time.

Contact details

If you have any problems you can contact us at any time on this telephone number: Dr Lindiwe Lamola, Dr Candice Feben or Prof Amanda Krause – 011 489 9223.

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

Hospital Street, Johannesburg, 2001 | PO Box 1038, Johannesburg, 2000
[T]: +27 11 489 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

PLEASE NOTE!

Some patients may experience psychological distress should a complication be detected. You can contact the Genetic Counselling Unit at the Division of Human Genetics, NHLS Braamfontein to arrange a session with an HPCSA-registered genetic counsellor to discuss these issues. They are experienced and qualified counsellors who are able to offer distress counselling, decision support, address concerns associated with genetic disorders and testing and explain medical terminology. Appointments can be arranged by contacting Ms Janet Robbie on (011) 489 9223 or Ms Shelley Macaulay on (011) 489 9243.

If you have any complaints or problems with this research you may report them to the Chair of the Human Research Ethics Committee at the University of the Witwatersrand on this telephone number – (011) 717 1234.

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

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 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

Protocol: M180855

Inherited cancer predisposing mutations in a cohort of childhood cancer patients

ASSENT FORM FOR CHILD PARTICIPANTS

PARTICIPANT STUDY ID:

The purpose of this study, procedures to be followed, risks and benefits, have been explained to me. I have been allowed to ask the questions I have, and my questions have been answered to my satisfaction. I have been told whom to contact if I have questions, to discuss problems, concerns, or suggestions related to the research, or to obtain information or offer input about the research. I have read this assent form (or had it read to me) and agree to participate in this research study with the understanding that I may withdraw at any time. I understand that that I may approach the researchers to be re-consented when I turn 18 if I want to and that I may withdraw from the study at any time. I have been told that I will be given a signed and dated copy of this consent form if I would like one.

Researcher	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Parent or Legal Guardian	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Child Participant	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Translator (if applicable)	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)

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

Information sheet – Child Participant

Inherited cancer predisposing mutations in a cohort of childhood cancer patients

Protocol: M180855

NATIONAL HEALTH LABORATORY SERVICE
UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

Division of Human Genetics
University of the Witwatersrand & National Health Laboratory Service
P.O. Box 1038
Braamfontein
Johannesburg
2000
Tel: +2711 489 9211
Fax : +2711 489 9226

Dear Potential Participant,

Hello, we are Dr. Lindie Lamola, Dr. Candice Feben, Prof. Amanda Krause and Dr Njabulo Mabaso from the Division of Human Genetics at the University of the Witwatersrand and National Health Laboratory Service. We are doing research on childhood cancers. We work with other members of the Division of Human Genetic to study the genetic factors that cause childhood cancers in South Africans.

What is cancer?



Cancer is an abnormal growth of cells. Cells are small structures which our bodies are made up of. Sometimes our cells get damaged and they start to grow differently and become big and abnormal – when they grow like this, they become what we call cancer.

What causes childhood cancers?

Cancer in children is difficult and sometimes doctors are not sure why some children get cancers and others don't.

For this study, we would like to look inside your DNA and find what is making your cells grow so much to cause cancer. By doing so, we will be able to understand how and why children like you have cancer.

What do I have to do?

Once you understand what the study is about and are happy to take part, we will ask you to sign a form to join the study. A copy of these forms will be given to you to keep.



Once you have become a participant in the study the doctor will do some checks like how heavy you are and how tall you are





The doctor will need to take some of your blood and/ or buccal swab which we will use to get your DNA.



Our work in this study is to test your DNA to find what caused your cancer



We will ask your Dr for permission to access your hospital file. This information will help us to understand your condition better

Long-term storing of samples and data



Your sample will be stored safely together with other participants' samples. The samples will only be used in the childhood cancer study and not in any other research.

Confidentiality

You will be given a code and number for the study so that no one knows your real name and who you are. All samples and the data will not be attached to your name.

You do not have to join this study. Whether you decide to join this study or not will not affect your treatment in our clinic.

What will I benefit from being part of this study?

You may not get any benefit directly from this study, but we hope it may help others who are affected by childhood cancers in the future. As it might help us to understand why you got cancer. We may contact you in future if we believe we have identified the cause of your condition.

Contact details



If you have any problems you can contact us at any time on this telephone number: Dr Lindie Lamola, Dr. Candice Feben, Dr Njabulo Mabaso or Prof Amanda Krause – 011 489 9223.



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 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

Protocol: M180855

Inherited cancer predisposing mutations in a cohort of childhood cancer patients

CONSENT FORM – PARENTAL CONSENT FOR CHILD PARTICIPANTS

PARTICIPANT STUDY ID

PERMISSION FOR PHOTOGRAPHS	Please initial
I agree to the use of deidentified photographs of my child for medical records ONLY.	
I agree for deidentified photographs my child to be used for teaching purposes AND to be used for medical records but NOT for medical publication.	
I agree for deidentified photographs of my child be used in medical publications, including medical journals, textbooks and electronic publications AND to be used for teaching purposes AND to be used for my child's medical records. I understand that the images may be seen by members of the general public, in addition to scientists and medical researchers that use these publications in their professional education. Although these photographs will be used without identifying information, such as my child's name, I understand that it is possible that someone may recognize my child.	

Researcher	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Parent Participant	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)

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 9211 | [F]: +27 11 489 9226 | [E]: human.genetics@nhls.ac.za

Protocol M180855

Inherited cancer predisposing mutations in a cohort of childhood cancer patients

CONSENT FORM – PARENTAL CONSENT FOR CHILD PARTICIPANTS

PARTICIPANT STUDY ID

PERMISSION FOR STORAGE		
I agree to have my samples stored in the research facility for future childhood cancer studies only without any identifiers.	Yes ☐	No ☐
I agree to have the data generated from my samples stored without any identifiers and be reanalysed if need be.	Yes ☐	No ☐

PERMISSION FOR REUSE				
	Samples		**Data**	
I agree that my child's samples & data (including clinical data) can be shared with other investigators for research *related to childhood cancers* once the ethics committee has approved the new study. My child's samples & data will be made available without any identifiers.	Yes ☐	No ☐	Yes ☐	No ☐

Researcher	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)
Parent Participant	☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐ Printed Name & Surname _____ ☐☐/☐☐/☐☐☐☐ Signature/Thumbprint Date (dd/mm/yyyy)

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

Appendix G:

Title Approval and Approved Research Protocol

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Private Bag 3 Wits, 2050
Fax: 027117172119
Tel: 02711 7172076

Reference: Mrs Sandra Benn
E-mail: sandra.benn@wits.ac.za

20 July 2023
Person No: 719425
PAG

Ms D Beukman
0C Montpark drive
Magalies unit 5
Randburg
2195
South Africa

Dear Ms Danielle Beukman

Master of Science in Medicine: Approval of Title

We have pleasure in advising that your proposal entitled *RB1 and beyond: Determining genetic causes of Retinoblastoma in South-African patients* has been approved. Please note that any amendments to this title have to be endorsed by the Faculty's higher degrees committee and formally approved.

Yours sincerely

Mrs Sandra Benn
Faculty Registrar
Faculty of Health Sciences

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

CANDIDATE'S SURNAME: **BEUKMAN**	FIRST NAME/S: **DANIELLE**	STUDENT NUMBER: **719425**
CURRENT QUALIFICATIONS: **BSc Hons Experimental Physiology**		

TEL: N/A	CELL: **0845834985**	E-MAIL: **beukmandanielle@gmail.com**	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: **2023**
DEPARTMENT: **HUMAN GENETICS**		
TITLE OF PROPOSED RESEARCH: **RB1 and beyond: Determining genetic causes of Retinoblastoma in South-African patients**		
CANDIDATE'S SIGNATURE: *BBeukman*		DATE: **7 July 2023**
SUPERVISOR 1 (NAME & SURNAME)**: Dr. Lindie Lamola**		50 % Supervision
SUPERVISOR'S QUALIFICATIONS **PhD**		
SUPERVISOR'S DEPARTMENT **Human Genetics**		
SUPERVISOR'S ADDRESS / TEL / E-MAIL: **lindie.lamola@wits.ac.za**		
SUPERVISOR 2 (NAME & SURNAME): **Nkateko Mayevu**		50 % Supervision
SUPERVISOR'S QUALIFICATIONS: **MSc**		
SUPERVISOR'S ADDRESS / TEL / E-MAIL: **nkateko.mayevu@wits.ac.za**		

SYNOPSIS OF RESEARCH:

Introduction:

Retinoblastoma (RB) is the most common intraocular malignancy in pediatric patients. Up to 31% of RB develops due to an inherited germline variant or mutation which predisposes an individual to RB development at a younger age. The most notable predisposing heritable mutation is found in the Retinoblastoma 1 (*RB1)* gene, however recent international genomics studies have identified other driver genes implicated in heritable RB.

Aim:

Identify causative germline variants predisposing pediatric South-African patients to RB.

Justification for study:

There is a lack of studies investigating mutations causing individuals with an African ancestry to retinoblastoma. From international studies, mutations in more than 43 genes have been proved to increase susceptibility of developing RB. The proposed study will determine if the South African RB patients also have likely pathogenic or pathogenic variants in these genes previously found to be associated with RB.

Proposed methodology:

This is a sub-study under the larger parent Pediatric Cancer (PeCan) project utilizing a cohort of eight pediatric RB patients recruited under the PeCan project. DNA from these patients will be used to perform whole exome sequencing (WES). The data obtained from sequencing will be analyzed using bioinformatics and in silico tools, to identify if they have germline variants in genes previously reported to be associated with RB development.

Expected results:

The majority of RB cases occur sporadically, heritable RB accounting for approximately 31% of RB cases. With a small cohort of eight patients, I may only find two to three germline variants. As with any genetic testing, if a germline variant is found it could be useful to family members of the patients.

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

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 Yes SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE: M180855 (M230782)
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

RB1 and beyond:

Determining genetic causes of retinoblastoma in South African patients

Danielle Beukman WITS

Student number 719425 MSc

(Med) Genomic Medicine

Supervisors:

Dr Lindie Lamola

Ms. Nkateko Mayevu





Table of Contents

1. INTRODUCTION 6

1.1 RETINOBLASTOMA PREVALENCE 6

1.2 GENETIC LANDSCAPE OF RETINOBLASTOMA 6

1.3 LOCUS HETEROGENEITY OBSERVED IN RETINOBLASTOMA 7

2. RATIONALE AND PURPOSE OF STUDY 10

3. AIM AND OBJECTIVES 10

3.1 AIM: 10

3.2 OBJECTIVES: 10

4. METHODS 11

4.1 WHOLE EXOME SEQUENCING WORKFLOW 11

4.2. BIOINFORMATICS ANALYSIS: 13

5. ETHICS 15

6. TIMING AND WORK PLAN 16

7. FUNDING AND BUDGET: 16

8. PROBLEMS AND LIMITATIONS 17

9. REFERENCES 18

10. APPENDIX I 23

1. INTRODUCTION

1.1 Retinoblastoma prevalence

Retinoblastoma (RB) is a solid intraocular cancer, most often developing in childhood. The South African National Cancer Registry (SANCR) recorded an average RB prevalence of 1 in 21 641 live births for the period of 2004 to 2018 (Stuart et al., 2022). The mean age of diagnosis in South Africa is 2.5 years, significantly later than mean age of diagnosis in high income countries, 1.2 years (Stuart et al., 2022). This is startling as the age of diagnosis is the strongest indicator of prognosis and whether an RB patient needs to undergo enucleation (Stuart et al., 2022).

1.2 Genetic landscape of retinoblastoma

The majority (65 - 68.2 %) of RB cases are classified as sporadic, occurring at a mean age of 24 months, presenting with unilateral tumors (Knudson, 1971). Heritable retinoblastoma accounts for the remaining 31.8% - 35% of RB cases, developing mostly bilateral tumors, with an earlier mean age of onset of 12 months (Knudson, 1971). In the case of heritable RB, an individual is predisposed to developing RB due to a likely pathogenic or pathogenic germline variant in the Retinoblastoma 1 gene. The first gene associated with RB predisposition was the Retinoblastoma 1 *(RB1)* tumor suppressor gene (TSG), discovered by Alfred Knudson (Knudson, 1971). To date more than 1700 germline pathogenic variants in *RB1* have been reported on the Leiden Open Variation Database (LOVD), which is an open-source database logging research-based DNA variation (Fokkema et al., 2011).

1.3 Locus heterogeneity observed in retinoblastoma

Next generation sequencing (NGS) such as whole exome sequencing (WES) of European and Chinese cohorts with heritable RB has identified pathogenic variants in genes other than *RB1* that contribute to RB development (Kooi et al., 2016). NGS performed on blood samples from patients with RB from Istanbul revealed this locus heterogeneity clearly. All patients reported a family history of RB, yet wild type *BR1* was observed in six individuals. When exome sequencing was performed on the six *RB1*-negative patients, twenty-seven significant variants (likely pathogenic or pathogenic) were found in genes acting in different biochemical pathways, such as retinoblast formation and DNA repair. Variants with the highest significance are found in the following genes: *FGFR4, NQO1, ACADS, KRT85, GBE1* and *CLEC7A* (Akdeniz et al., 2019).

Mutations in the TSG *p53* have long been associated with heritable cancer syndromes. In a large study with 111 retinoblastoma patients with a family history of RB of Italian ancestry, single nucleotide variants or mutations were found in *p53* and *MDM2* (Epistolato et al., 2011). The TSG *p16INK4A* is involved in cellular regulation preventing inappropriate cellular proliferation, halting the cell cycle at the G1-S checkpoint by encoding a CDK inhibitor (Epistolato et al., 2011). In a study of 29 Italian patients with RB, promoter methylation and downregulation of *p16INK4A* was strongly associated with heritable RB but not with sporadic RB (Indovina et al., 2009). Confirming the heritable aspect of an epigenetic susceptibility marker and reaffirmed the significance of *p16INK4A* in heritable RB (Indovina et al., 2009).

Together MDM2 and MDM4 regulate p53, due to the interaction in a common biochemical pathway. *MDM2* was confirmed to be a modifier gene in families with

RB malignancies (Castéra et al., 2010). Single nucleotide variants in both *MDM2* and *MDM4* were strongly implicated in heritable RB. This study also found that the *MDM4* rs116197192G allele frequency was significantly higher in their Brazilian RB cohort compared to controls (de Oliveira Reis et al., 2012).

A study by Carvalho and colleagues from 2013 investigated samples from 141 RB patients and 120 controls, this study revealed pathogenic variants in *CDKN1A* gene in the presence of wild type *RB1* gene. *CDKN1A*, when mutated, contributes to uncontrolled cell proliferation (Carvalho et al., 2013).

PAX6 gene encodes the PAX6 transcription factor that regulates retinogenesis. Inactivation or suppression of *PAX6* provides retinoblasts a molecular mechanism to obtain two cancer hallmarks, resisting cell death by inhibiting apoptosis and sustained proliferative signaling. Pathogenic or likely pathogenic germline variants in *PAX6* have been reported in RB patients with wildtype *RB1* (Meng et al., 2014). In this instance *PAX6* acts as a TSG, yet in other studies *PAX6* was overexpressed in pancreatic tumors, indicating that PAX6 can be classified as an oncogene (lang et al., 2008). The molecular duality of the *PAX6* gene complicates the classification of variants but is believed to differ according to the type of cancer (Lang et al., 2008).

Pathogenic germline variants in oncogenes have been found to be major driver genes in RB development (Afshar et al., 2020). Likely pathogenic or pathogenic germline variants were reported in the following oncogenes: *MYCN* and *TSC2* (Francis et al., 2021).

The genes implicated in RB malignancy are summarized in Table 1 below.

Table 1: Genes implicated in RB development.

Gene Symbol	Full gene name	Reference
RB1	Retinoblastoma	Knudson, 1971
ACADS	Short-chain acyl-CoA dehydrogenase	Akdenizet al., 2019
CLEC7A	C-Type Lectin Domain Containing 7A	Akdenizet al., 2019
CDK	Cyclin-dependent kinases	Epistolato et al., 2011
CDKN1A	Cyclin Dependent Kinase Inhibitor 1A	Carvalho et al., 2013
FGFR4	Fibroblast growth factor receptor 4	Akdenizet al., 2019
GBE1	Glycogen branching enzyme	Akdenizet al., 2019
KRT85	Keratin 85	Akdenizet al., 2019
MDM2	E3 ubiquitin-protein ligase	Epistolato et al., 2011
MDM4	P53-Binding Protein	de Oliveira Reis et al., 2012
MYCN	MYCN Proto-Oncogene, also called BHLH Transcription Factor	Francis et al., 2021
NQO1	Dehydrogenase [quinone] 1	Akdenizet al., 2019
PAX6	Aniridia type II	Lang et al., 2008 Meng et al., 2014
p16INK4A	Cyclin-dependent kinase inhibitor 2A	Indovina et al., 2009
p53	TP53	Epistolato et al., 2011
TSC2	Tuberous Sclerosis Complex 2	Francis et al., 2021

2. RATIONALE AND PURPOSE OF STUDY

Currently there is a great gap in the literature with regards to genetic variants contributing to pediatric cancer development in patients with a South African ancestry, this includes RB. Studies performed in developed countries have found numerous candidate genes that potentially contributing to an individual's likelihood of developing hereditary RB, similar studies have not been conducted in Africa. The purpose of this study is to determine if the South African cohort of RB patients have likely pathogenic or pathogenic variants in genes previously reported to be associated with RB.

3. AIM AND OBJECTIVES

3.1 AIM:

Identify causative germline variants implicated in RB development and progression in pediatric South-African patients.

3.2 OBJECTIVES:

- Perform whole exome sequencing (WES) on four DNA samples from South- African patients affected with RB.
- Perform bioinformatics analysis of the exome sequence data from 8 patients with RB to identify candidate likely pathogenic and pathogenic variants associated with RB development and progression. The genes to be analyzed are listed in Table 1.

4. METHODS

Eight pediatric patients diagnosed with RB were recruited as part of the larger parent study, Pediatric Cancer Project (PeCan). All patient's personal information accessed as part of this sub-study will be anonymized using patient codes. Patient information is encrypted, and password protected. Patient history includes race, age of diagnosis, sex, and tumor presentation. Patient information is saved on the REDCap database, accessible to researchers, I have been granted access by the gatekeeper. Clinical data will be used in conjunction with WES data as part of result interpretation. DNA was isolated from peripheral blood samples. Of the eight patients, WES has been performed on four samples (1-4), for the remaining four samples (5-8) WES will be performed as a step in this project.

4.1 Whole Exome Sequencing Workflow

The stepwise WES workflow is summarized in figure 1.

4.1.1. Assess DNA Quality:

Initial DNA quantification will be performed using the Thermo Scientific Nanodrop Spectrophotometer ™ to assess DNA concentration, this densitometric method will be followed by using The Thermo Scientific Qubit ™ fluorometer to assess double stranded DNA concentration, as this will be the starting material in WES.

4.1.2. Library Preparation and Amplification:

Library preparation and amplification will be conducted as per manufacturer's instructions provided in the Ion AmpliSeq™ manuals (ThermoFisher Scientific, Waltham MA, USA). Library preparation entails target amplification, amplicon partial digest and ligation of IonCode™ Barcode adapters. At this stage the unamplified library will be purified. Subsequently, purified libraries will be amplified and further purified.

4.1.3. Library Quantification and dilution:

Quantification of the amplified library will be performed using a Qubit™ 3 fluorometer (ThermoFisher Scientific, Waltham MA, USA). This will be followed by Tape Station™ (Agilent technologies, Santa Clara, USA) which will quantify the libraries and determine concentration based on fragment sizes of interest, between 200- 400 bp. This allows for the appropriate dilutions to be made prior to templating.

4.1.4. Templating and Sequencing:

Templating of the libraries occur in a flow cell, the ion sphere particles contain primer sequences that will bind to complementary single stranded DNA, this is facilitated by the barcode adapters. Templating creates a sphere covered with multiple copies of the same amplicon, ensuring that the signals produced during sequencing are strong enough to be detected. Templating will be performed using the Ion 540™ Kit – Chef (ThermoFisher Scientific, Waltham MA, USA), followed by exome sequencing performed in by the Ion torrent S5 semiconductor™ sequencer machine (ThermoFisher Scientific, Waltham MA, USA).

This proposed whole exome sequencing workflow is illustrated in Figure 1.

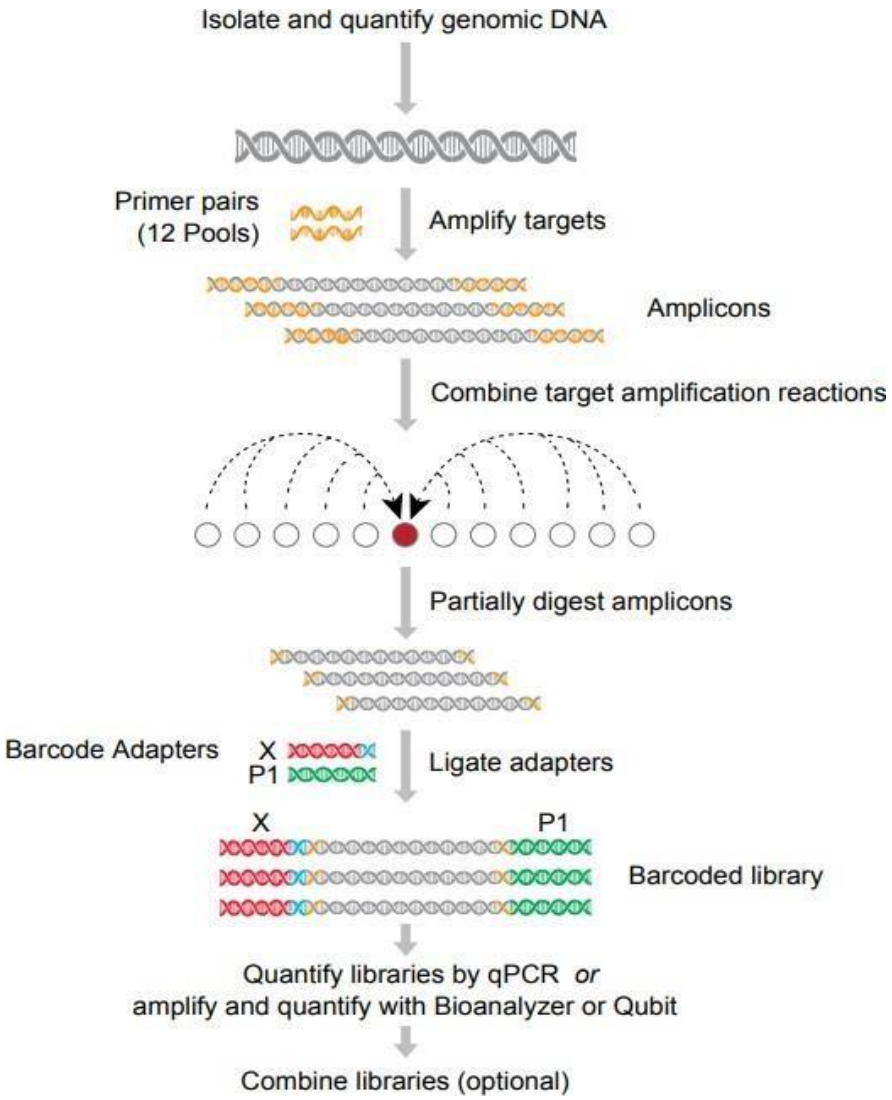


Figure 1: Ion AmpliSeq™ Exome RDY library preparation (ThermoFisher Scientific manual MAN0010084 version E.0, 2018). Xand P1 represent the barcoded adapters that get ligated to make each amplicon unique and identifiable.

4.2. Bioinformatics Analysis:

Once all raw data has been generated from whole exome sequencing, bioinformatic analysis will be employed to search for variants of interest. WES raw data is generated as a variant call file (VCF file) by aligning the sequences in question to a human reference genome

GRCh37/Hg19 and identifying sequence differences between the two. The GRCh37/Hg19 reference genome will be used as reference genome as the reagents (Ion AmpliSeq ™ Exome RDY kit) in stock contains primers designed according to the GRCh37/Hg19 reference genome. The sequencer instrument (Ion GeneStudio ™ S5 Instrument) software currently in use by the division utilizes analysis pipelines created for GRCh27/Hg19. A variant effect predictor (VEP) from Ensamble Genome Browser will be used to annotate the variants found in genes and transcripts (McLaren et al., 2016). Annotation output will include the following results: Variant location (chromosome, gene, and exon) and the type of variant detected (i.e., missense, nonsense, and frameshift) (Hunt et al., 2021). The frequency of the variants will be ascertained from two genomics databases, 1000 genomes project (The 1000 Genomes Project Consortium et al., 2015) and gnomAD (Gudmundsson et al., 2022).

Variant prioritization is essential to reduce the large number of variants generated and identify possible disease-causing variants. First line of variant prioritization involves selecting genes previously reported to be involved in RB development, as mentioned in Table 1. Variants with a minor allele frequency higher than 5% will be filtered out (this figure is subject to change depending on the number of results that get returned) as well as variants that have no effect on protein function or structure as predicted by in silico prediction tools such as Sort Intolerable from tolerable (SIFT) (Ng & Henikoff, 2003) and Polymorphism Phenotyping version 2 (POLYPHEN-2) (Adzhubei et al., 2010). These in silico tools predict what structural and functional change will occur in a protein due to an amino acid change (de la Campa et al. 2017).If the ClinVar database has available data on the variants of interest, the variant and gene-associated phenotype will be assessed, to aid in filtering out variants that have been reported as benign or likely benign.

A Bam.file and the shortlisted variants will be viewed in an integrative genomics viewer (IGV) to assess the quality of variant calling and coverage. The remaining shortlisted variants will be classified according to The American College of Medical Genetics and Genomics and Association for Molecular Pathology Guidelines (ACMG/AMP) which aggregates variant data from different predictors and research-based findings into a score (Richards et al., 2015).

All possible causative variants will be reported, these include variants of uncertain significance, likely pathogenic variants, and pathogenic variants. If a reportable variant is found, this information will be passed on to the Principal Investigator of the PeCan project parent study, Dr. Lindie Lamola. As part of the parent study these variants will be validated. Information will be shared with the relevant clinicians and genetic counselors to communicate findings in accordance with the ethical permissions obtained at the time of sample collection.

5. ETHICS

Ethics approval for the parent study, has been acquired (M180855). Ethical clearance for this project has been approved conditionally by the Wits Human Research Ethics Committee (Wits HREC) pending protocol approval by the Faculty of Health Sciences. Parent study ethics study is attached in Appendix I.

6. TIMING AND WORK PLAN

Study to commence as soon as ethics clearance is granted. Study to be completed by 30 Nov 2023.

	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Literature review										
Ethics application										
Preparing protocol										
Protocol assessment										
NGS Lab (WES)										
Data Analysis										
Paper write-up										

7. FUNDING AND BUDGET:

All costs of this study will be covered by funding already obtained for the parent study,The PeCan Project and is managed by the principal investigator Dr. Lindie Lamola. Funding has been obtained from the following partners: National Research Foundation (NRF), South African Medical Research Council Self-Initiated Research (SAMRC SIR), and Female Academic Leadership Fellowship (FALF).

Table 2: Exome sequencing budget

Exome sequencing	
ION AMPLISEQ EXOME RDY KIT	R16775
ION 540 CHIP KIT EACH x2	R 25375
ION 540 KIT-CHEF (2/INIT)	R28900
DNA and library quantification	
Qubit reagents	R1500
Tape station reagents	R2800
Total	R 75350

8. PROBLEMS AND LIMITATIONS

A small sample size of 8 RB patients increases the likelihood of no significant variants being found. Short read sequence data will be obtained which may lead to alignment issues. Due to the lack of African genomics data and European bias, variants common in African populations but rare in European populations might be misinterpreted as likely pathogenic or pathogenic. The in-silico predictors that will be utilized are notorious for reporting contradictory and conflicting results. We are not using the latest reference genome because at the time of sequencing the first four exomes, Hg19 was used as the reference genome.

9. REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Afshar, A. R., Pekmezci, M., Bloomer, M. M., Cadenas, N. J., Stevers, M., Banerjee, A., Roy,R., Olshen, A. B., Van Ziffle, J., Onodera, C., Devine, W. P., Grenert, J. P., Bastian, B. C., Solomon, D. A., & Damato, B. E. (2020). Next-Generation Sequencing of Retinoblastoma Identifies Pathogenic Alterations beyond RB1 Inactivation That Correlate with Aggressive Histopathologic Features. *Ophthalmology*, *127*(6), 804–813.https://doi.org/10.1016/j.ophtha.2019.12.005

Akdeniz, D., Tuncer, S. B., Kebudi, R., Celik, B., Kuru, G., Kilic, S., Sukruoglu Erdogan, O., Avsar, M., Buyukkapu Bay, S., Tuncer, S., & Yazici, H. (2019). Investigation of new candidate genes in retinoblastoma using the TruSight One “clinical exome” gene panel. *Molecular Genetics & Genomic Medicine*, *7*(8), 1-14. https://doi.org/10.1002/mgg3.785

Carvalho, I. N. S. R., Reis, A. H. de O., Cabello, P. H., & Vargas, F. R. (2013). Polymorphisms of CDKN1A gene and risk of retinoblastoma. *Carcinogenesis*,*34*(12),2774–2777.https://doi.org/10.1093/carcin/bgt308

Castéra, L., Sabbagh, A., Dehainault, C., Michaux, D., Mansuet-Lupo, A., Patillon, B., Lamar,E., Aerts, I., Lumbroso-Le Rouic, L., Couturier, J., Stoppa-Lyonnet, D., Gauthier- Villars, M., & Houdayer, C. (2010). MDM2 as a Modifier Gene in Retinoblastoma. *JNCI:Journal of the National Cancer Institute*, *102*(23), 1805– 1808. https://doi.org/10.1093/jnci/djq416

de la Campa, E. álvarez, Padilla, N., & de la Cruz, X. (2017). Development of

pathogenicity predictors specific for variants that do not comply with clinical guidelines for the use of computational evidence. BMC Genomics, 18(Suppl 5), 1–14. https://doi.org/10.1186/s12864-017-3914-0

de Oliveira Reis, A. H., de Carvalho, I. N. S. R., de Sousa Damasceno, P. B., Ferman, S. E.,Lucena, E., Lopez-Camelo, J. S., Seuánez, H. N., & Vargas, F. R. (2012). Influence of MDM2 and MDM4 on development and survival in hereditary retinoblastoma. *Pediatric Blood & Cancer*, *59*(1), 39–43. https://doi.org/10.1002/pbc.24014

Epistolato, M. C., Disciglio, V., Livide, G., Berchialla, P., Mencarelli, M. A., Marozza, A., Amenduni, M., Hadjistilianou, T., De Francesco, S., Acquaviva, A., Toti, P., Cetta, F., Ariani, F., De Marchi, M., Renieri, A., & Giachino, D. (2011). P53 Arg72Pro and MDM2309 SNPs in hereditary retinoblastoma. *Journal of Human Genetics*, *56*(9), 685–686. https://doi.org/10.1038/jhg.2011.82

Fokkema, I. F., Taschner, P. E., Schaafsma, G. C., Celli, J., Laros, J. F., & den Dunnen, J. T. (2011). LOVD v.2.0: the next generation in gene variant databases. Humanmutation, 32(5),557–563. https://doi.org/10.1002/humu.21438

Francis, J. H., Richards, A. L., Mandelker, D. L., Berger, M. F., Walsh, M. F., Dunkel, I. J., Donoghue, M. T. A., & Abramson, D. H. (2021). Molecular Changes in Retinoblastoma beyond RB1: Findings from Next-Generation Sequencing. *Cancers*, *13*(1), 149 112.https://doi.org/10.3390/cancers13010149

Gudmundsson, S., Singer-Berk, M., Watts, N. A., Phu, W., Goodrich, J. K., Solomonson, M., Genome Aggregation Database Consortium, Rehm, H. L., MacArthur, D. G., & O'Donnell-Luria, A. (2022). Variant interpretation using population databases: Lessons from gnomAd. *Human mutation*, *43*(8),1012–1030 https://doi.org/10.1002/humu.24309

Hunt, S. E., Moore, B., Amode, R. M., Armean, I. M., Lemos, D., Mushtaq, A., Parton, A., Schuilenburg, H., Szpak, M., Thormann, A., Perry, E., Trevanion, S. J., Flicek, P., Yates, A. D., & Cunningham, F. (2021). Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor—A tutorial. *Human Mutation*, 1–12. https://doi.org/10.1002/humu.24298

Indovina, P., Acquaviva, A., De Falco, G., Rizzo, V., Onnis, A., Luzzi, A., Giorgi, F., Hadjistilianou, T., Toti, P., Tomei, V., Pentimalli, F., Carugi, A., & Giordano, A. (2009). Downregulation and aberrant promoter methylation of *p16INK4A*: A possible novel heritable susceptibility marker to retinoblastoma. *Journal of Cellular Physiology*, 143-150. https://doi.org/10.1002/jcp.22019

Kooi, I. E., Mol, B. M., Massink, M. P. G., Ameziane, N., Meijers-Heijboer, H., Dommering, C. J., van Mil, S. E., de Vries, Y., van der Hout, A. H., Kaspers, G. J. L., Moll, A. C., te Riele, H., Cloos, J., & Dorsman, J. C. (2016). Somatic genomic alterations in retinoblastoma beyond RB1 are rare and limited to copy number changes. *Scientific Reports*, *6*(1), 25264. https://doi.org/10.1038/srep25264

J. C. (2016). A Meta-Analysis of Retinoblastoma Copy Numbers Refines the List of Possible Driver Genes Involved in Tumor Progression. *PLOS ONE*, *11*(4), e0153323,1 - 20. https://doi.org/10.1371/journal.pone.0153323

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R. S., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122. https://doi.org/10.1186/s13059-016-0974-4

Meng, B., Wang, Y., & Li, B. (2014). Suppression of PAX6 promotes cell proliferation and inhibits apoptosis in human retinoblastoma cells. *International Journal of Molecular Medicine*, *34*(2), 399–408. https://doi.org/10.3892/ijmm.2014.1812

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic-acids-research*, *31*(13),3812–3814. https://doi.org/10.1093/nar/gkg509

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. (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), 405–424. https://doi.org/10.1038/gim.2015.30

Stuart, K. V., Shepherd, D. J., Kruger, M., & Singh, E. (2022). The Incidence of Retinoblastoma in South Africa: Findings from the South African National Cancer Registry (2004–2018).*Ophthalmic-Epidemiology*,*-29*(6),681–687. https://doi.org/10.1080/09286586.2021.2013900

The 1000 Genomes Project Consortium, Corresponding authors, Auton, A., Abecasis, G. R., Steering committee, Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., Donnelly, P., Eichler, E. E., Flicek, P., Gabriel, S. B., Gibbs, R. A., Green, E. D., Hurles, M. E., Knoppers, B. M., … Abecasis, G. R. (2015). A global reference for human genetic variation. *Nature*, *526*(7571), 68–74. https://doi.org/10.1038/nature15393

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., … Downing, J. R. (2015). Germline Mutations in Predisposition Genes in Pediatric Cancer. *New England Journal of Medicine*, *373*(24), 2336–2346. https://doi.org/10.1056/NEJMoa1508054

10. Appendix I

R14/49 Dr L Lamola, et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180855

NAME: (Principal Investigator)	Dr L Lamola, et al
DEPARTMENT:	School of Pathology Department of Human Genetics National Health Laboratory Service
PROJECT TITLE:	Inherited cancer predisposing mutations in a cohort of childhood cancer patients
DATE CONSIDERED:	31/08/2018
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Not applicable
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	26/02/2019

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

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

I/We fully understand the conditions under which I am/we are authorised 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 resubmit to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date of the meeting when the study was initially reviewed. In this case, the study was initially reviewed in **August** and will therefore reports and re-certification will be due early in the month of **August** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

26/02/2019
Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix H:

Plagiarism declaration and Turnitin report

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 Danielle Beukman (Student number: 719425) am a student registered for the degree of MSc Med Genomic Medicine in the academic year 2023.

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:  Date: 27 November 2023

turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Danielle Beukman
Assignment title:	Research Report
Submission title:	DBeukman Research report (Paper Sub format) .pdf
File name:	DBeukman_Research_report_28Paper_Sub_format_29_.pdf
File size:	412.02K
Page count:	17
Word count:	5,873
Character count:	32,261
Submission date:	27-Nov-2023 10:21PM (UTC+0200)
Submission ID:	2240005394



1. Abstract for Journal Submission

Background: Retinoblastoma is the most common solid tumour of the retina, affecting mainly paediatric patients. Knudson formulated the "two-hit" hypothesis on the grounds of the Retinoblastoma 1 (*RB1*) gene, the first tumour suppressor gene discovered, laying the foundation to investigate how germline variation in cancer predisposition genes can genetically predispose an individual to malignancy. Although loss-of-function germline variants the *RB1* gene is the tumour suppressor gene most often associated with heritable retinoblastoma, two confounding research findings complicate the assumption that a pathogenic germline variant in *RB1* will lead to the development of heritable retinoblastoma. First, heritable retinoblastoma has been observed in cohorts with biallelic wild-type *RB1* genes, secondly, biallelic loss-of-function is not always sufficient to lead to retinoblastoma tumorigenesis. The "beyond *RB1*" concept addresses these contradictions by investigating germline variants in additional cancer predisposition genes to discover candidate driver genes in heritable retinoblastoma. In this study we investigate sequence germline variants in cancer predisposition genes beyond *RB1* in retinoblastoma patients with a South African ancestry.

Methods: Peripheral blood samples from eight paediatric patients of South African ancestry diagnosed with retinoblastoma was used to extract genomic DNA. Whole exome sequencing with single nucleotide resolution was performed to establish germline sequence variants. Variant annotation, prioritisation and filtering was executed to detect germline variants in cancer predisposition genes previously proposed as candidate genes in retinoblastoma malignancy and paediatric cancer predisposition syndromes. Variants shared between individuals and variants of interest were classified according to ACMG/AMP guidelines.

Results: A total of 353 655 variant were annotated, 4538 of which were novel. After collating functional data, previously reported and reviewed *RB1* variants, all sequence variants in *RB1* are likely benign, amending variant classification based on updated in-silico predictors. The extended list of cancer predisposing genes revealed variants of uncertain significance.

Conclusion: The absence of sequence variants capable of describing the retinoblastoma phenotype in this cohort demonstrates the necessity of combining next generation sequencing with analysis pipelines and additional genome mapping to detect structural variants. This study adds genomic information from patients with a South African ancestry to mounting genomic data, ensuring that previously underrepresented populations can also benefit from future cancer predisposition and precision medicine research.

Keywords: genetic predisposition, germline variation, cancer predisposition genes, paediatric tumours, whole exome sequencing, *RB1* gene, retinoblastoma, sequence variants

Copyright 2023 Turnitin. All rights reserved.

DBeukman Research report (Paper Sub format) .pdf

by Danielle Beukman

Submission date: 27-Nov-2023 10:21PM (UTC+0200)
Submission ID: 2240005394
File name: DBeukman_Research_report_28Paper_Sub_format_29_.pdf (412.02K)
Word count: 5873
Character count: 32261

DBeukman Research report (Paper Sub format) .pdf

ORIGINALITY REPORT

9%	6%	6%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet Source	1%
2	iris.unige.it Internet Source	1%
3	Lina Basel-Salmon, Noa Ruhrman-Shahar, Naama Orenstein, Michal Levy et al. "Clinically actionable incidental and secondary parental genomic findings after proband exome sequencing: yield and dilemmas", Genetics in Medicine Open, 2023 Publication	1%
4	Submitted to University of Witwatersrand Student Paper	<1%
5	www.mdpi.com Internet Source	<1%
6	academic.oup.com Internet Source	<1%
7	Elisabeth Jarhelle, Hilde Monica Frostad Riise Stensland, Geir Åsmund Myge Hansen, Siri Skarsfjord et al. "Identifying sequence	<1%

variants contributing to hereditary breast and ovarian cancer inBRCA1 and BRCA2 negative breast and ovarian cancer patients", Scientific Reports, 2019
Publication

8	onlinelibrary.wiley.com Internet Source	<1 %
9	www.ajol.info Internet Source	<1 %
10	jasn.asnjournals.org Internet Source	<1 %
11	J. Weiner. "Evolution of Circular Permutations in Multidomain Proteins", Molecular Biology and Evolution, 02/01/2006 Publication	<1 %
12	Waad Rashed Al-Dosari, Ibrahim Abdullah Al-Binali, Satya Sudheer Pydi, William Goodwin. "Analysis of forensic casework samples by Precision ID Ancestry panel – manual and automated Ampliseq workflow", Forensic Science International: Genetics Supplement Series, 2019 Publication	<1 %
13	www.frontiersin.org Internet Source	<1 %
14	Calì, Francesco, Valeria Chiavetta, Giuseppa Ruggeri, Maria Piccione, Angelo Selicorni,	<1 %

Daniela Palazzo, Maria Bonsignore, Anna Cereda, Maurizio Elia, Pinella Failla, Maria Grazia Figura, Agata Fiumara, Silvia Maitz, Giuseppa Maria Luana Mandarà, Teresa Mattina, Alda Ragalmuto, Corrado Romano, Martino Ruggieri, Roberto Salluzzo, Antonino Saporoso, Carmelo Schepis, Giovanni Sorge, Maria Spanò, Gaetano Tortorella, and Valentino Romano. "Mutation spectrum of NF1 gene in Italian patients with neurofibromatosis type 1 using Ion Torrent PGM™ platform", European Journal of Medical Genetics, 2016.
Publication

15	Lorena Alonso, Anthony Piron, Ignasi Morán, Marta Guindo-Martínez et al. "TIGER: The gene expression regulatory variation landscape of human pancreatic islets", Cell Reports, 2021 Publication	<1 %
16	www.gsea-msigdb.org Internet Source	<1 %
17	Se Kyung Lee, Wan Wook Kim, Sung Hoon Kim, Sung Mo Hur et al. "Characteristics of metastasis in the breast from extramammary malignancies", Journal of Surgical Oncology, 2010 Publication	<1 %

18	genomemedicine.biomedcentral.com Internet Source	<1 %
19	molecularcasestudies.cshlp.org Internet Source	<1 %
20	Ana Rita Simões, Ceres Fernández-Rozadilla, Olalla Maroñas, Ángel Carracedo. "The Road so Far in Colorectal Cancer Pharmacogenomics: Are We Closer to Individualised Treatment?", Journal of Personalized Medicine, 2020 Publication	<1 %
21	Dewald Eygelaar, Elizabeth J. van Rensburg, Fourie Joubert. "Germline sequence variants contributing to cancer susceptibility in South African breast cancer patients of African ancestry", Scientific Reports, 2022 Publication	<1 %
22	Toma,. "Molecular Aspects of Transforming Growth Factor beta (TGF-beta) Resistance in Atherosclerosis: Role of Smad3 in Antiproliferative Response", Proquest, 2012. Publication	<1 %
23	haematologica.org Internet Source	<1 %
24	rcastoragev2.blob.core.windows.net Internet Source	<1 %

25	repository.ubn.ru.nl Internet Source	<1 %
26	respiratory-research.biomedcentral.com Internet Source	<1 %
27	storage.googleapis.com Internet Source	<1 %
28	www.biotnt.com Internet Source	<1 %
29	www.medrxiv.org Internet Source	<1 %
30	Jia Yu, Pei-Ju Liao, Weijun Xu, Julie R. Jones, David B. Everman, Heather Flanagan-Steet, Thomas H. Keller, David M. Virshup. "Structural model of human PORCN illuminates disease-associated variants and drug-binding sites", Journal of Cell Science, 2021 Publication	<1 %
31	Rashmi Chowdhary. "Interrupting wild poliovirus transmission using oral poliovirus vaccine: Environmental surveillance in high-risks area of India", Journal of Medical Virology, 08/2008 Publication	<1 %
32	Wan Nur Baitty Wan Mohd Tajuddin, Faridah Abas, Iekhsan Othman, Rakesh Naidu.	<1 %

"Molecular Mechanisms of Antiproliferative and Apoptosis Activity by 1,5-Bis(4-Hydroxy-3-Methoxyphenyl)1,4-Pentadiene-3-one (MS13) on Human Non-Small Cell Lung Cancer Cells", International Journal of Molecular Sciences, 2021

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

Appendix I:

Journal Author Guidelines for Submissible format

Please find the following link to access the author guidelines for the Journal Human Mutation.

https://www.hindawi.com/journals/humu/guidelines/