

https://dspace.org/

Identification of Pathogenic Variants in African Individuals with Oculocutaneous Albinism Type 2 Negative for the Common 2.7 kb Deletion Variant

Moodley, Priya

2025
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/48241

Moodley, Priya. (2025). Identification of Pathogenic Variants in African Individuals with Oculocutaneous Albinism Type 2 Negative for the Common 2.7 kb Deletion Variant [Master`s dissertation, University of the Witwatersrand, Johannesburg]. WIReDSpace. https://hdl.handle.net/10539/48241

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

IDENTIFICATION OF PATHOGENIC VARIANTS IN AFRICAN INDIVIDUALS WITH OCULOCUTANEOUS ALBINISM TYPE 2 NEGATIVE FOR THE COMMON 2.7 KB DELETION VARIANT

Priya Moodley

Student number: 2111913

Supervisor: Dr Thandiswa Ngcungcu

Co-supervisors: Dr Robyn Kerr and Dr Phelelani T. Mpangase



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

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

Johannesburg, 2025

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

Re: Research Report submitted to the Faculty of Health Sciences by Ms. Priya Moodley in partial fulfilment of the requirements for the degree of Master of Science in Medicine (Genomic Medicine)

Dear examiner,

Thank you for agreeing to examine the research report titled "*Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant*".This research report is being submitted by the candidate (Ms. Priya Moodley, student number: 2111913) in the format of a "submissible" paper as per the style guide for Thesis, Dissertation, and Research Reports (Updated March 2016) by the University of Witwatersrand, Faculty of Health Sciences. This Research paper contributes 33% towards the final mark of the MSc (Med): Genomic Medicine degree forwhich the candidate is enrolled.

The candidate and her supervisors, Dr. Thandiswa Ngcungcu, Dr. Robyn Kerr and Dr. Phelelani T. Mpangase, have chosen to submit the Research Report to the Frontiers in Genetics journal, therefore the manuscript was formatted according to the author guidelines of this journal. This international journal is a peer-reviewed, open-access journal published by Frontiers Media. The author guidelines have been attached for your reference (Appendix G). There are deviations made from the journal submission guidelines (e.g. tables and figures are placed in text, line numbers are not inserted, line spacing of 1.5 was used and manuscript is presented as one complete document) for ease of reading and marking.

The Research Protocol was assessed and approved by the Faculty of Health Sciences Post Graduate Assessors Committee and is attached below (Appendix D) for a detailed literature review and to further contextualise the study but is not for examination purposes.

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

Yours sincerely

Candidate:
Priya Moodley

Supervisor:
Dr. Thandiswa Ngcungcu

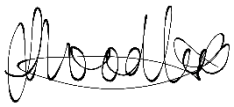
Co-Supervisor:
Dr. Robyn Kerr

Co-Supervisor:
Dr. Phelelani T. Mpangase

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

DECLARATION

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



Priya Moodley

Student number: 2111913

3rd day of March 2025 in Johannesburg

CONTRIBUTION OF CANDIDATE

Student's contribution to article and agreement of co-authors

I, Priya Moodley, declare that this Research Report entitled, "Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant", is my own work and that I contributed significantly towards the research findings presented in the paper intended for publication below.



26/11/24

Signature of candidate: **Date**

Priya Moodley



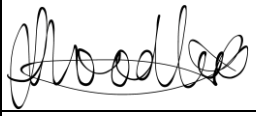
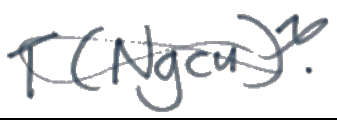
27/11/24

Signature of supervisor: **Date**

Dr Thandiswa Ngcungcu

Agreement by co-authors:

By signing this declaration, the co-authors listed below agree to the use of the article bythe candidate as part of her Research Report.

Authors	**Name**	**Signature**	**Date**
1st author	Priya Moodley		26/11/24
2nd author	Dr Robyn Kerr		26/11/24
3rd author	Dr Phelelani T Mpangase		26/11/24
4th author	Dr Thandiswa Ngcungcu		27/11/24

DEDICATION

To my parents, Lovashini and Sevalen, whose unwavering love and enduring values have guided and inspired me throughout this journey—this achievement is because of you.

PRESENTATIONS ARISING FROM THIS RESEARCH

1. Young and Early Career Geneticists Group (YEGG) Symposium, 26th October 2024 (poster presentation)
2. Southern African Society for Human Genetics (SASHG) 2024 Biennial Congress, 27th – 29th October 2024 (poster presentation)

Abstract

Oculocutaneous albinism is a group of rare, inherited pigmentation disorders characterised by hypopigmentation of the hair and skin and impaired vision. The most common form of albinism is oculocutaneous albinism type 2 (OCA2) which has a high prevalence in African populations. The most common cause of OCA2 in African individuals is a 2.7 kb intragenic deletion variant in the *OCA2* gene, and in the Southern African population, it accounts for 78% of all *OCA2* pathogenic variants. The aim of this study was to identify pathogenic variants in African individuals with OCA2 that were negative or heterozygous for the common African *OCA2* 2.7 kb deletion. Whole exome sequencing (WES) data was available for eight individuals who were clinically diagnosed with OCA2 and referred to the Molecular Genetics Diagnostic Laboratory, Division of Human Genetics, NHLS for genetic testing. The WES data was analysed to identify possible pathogenic variants in the *OCA2* gene, besides the 2.7kb deletion, and variants in other OCA-related genes (where no variants were identified in the *OCA2* gene). Variants were annotated using Ensembl's Variant Effect Predictor and then filtered based on functionality and minor allele frequency. Frameshift, missense, non-synonymous and splice site variants with a low minor allele frequency (MAF<5%) were prioritised. Prioritised variants were classified according to the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guidelines. This study identified a likely pathogenic variant, NM_000275.3: c.1239+2T>A, in the *OCA2* gene and a variant of uncertain significance, NM_000550.3: c. 291G>C, in the *TYRP1* gene in separate patients. These results suggest that WES is a useful technique for identifying variants in individuals with albinism, a disorder that exhibits locus heterogeneity. Limitations regarding copy number analysis using WES data, and difficulties encountered in interpretation of variants due to lack of population data and functional studies, should be addressed in future research. Identification of these potentially clinically significant variants contributes to our understanding of the variant profile of the OCA2 phenotype in individuals of African descent.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisors, Dr Thandiswa Ngcungcu, Dr Phelelani Mpangase, and Dr Robyn Kerr for their invaluable guidance, mentorship, and encouragement throughout this project. Your expertise and insights have been instrumental in shaping the direction and outcomes of my research on oculocutaneous albinism type 2. I am particularly grateful for your constructive feedback and unwavering support, which motivated me to strive for excellence. To Dr Thandiswa Ngcungcu, thank you for your continued support throughout my entire postgraduate journey. Thank you for always encouraging me to excel, trusting me and my work and inspiring me to make a meaningful contribution to the world of human genetics. To Dr Phelelani T. Mpangase, thank you for your incredible expertise in bioinformatics and for your patience and guidance in teaching me how to navigate the world of coding. To Dr Robyn Kerr, thank you for support and guidance not only for this project, but also for my academic journey and future endeavours. I appreciate the commitment that all three of you gave to fostering a collaborative and enriching research environment, which made this journey both challenging and rewarding.

My gratitude extends to the University of the Witwatersrand and the Kirsch Foundation for awarding me the Wits Postgraduate Scholarship. This funding opportunity made it possible for me to complete this degree. And a special thanks again to Dr Thandiswa Ngcungcu for funding this project, under your parent project, through your MRC-SIR grant.

Thank you to everyone at the Division of Human Genetics at the NHLS, Braamfontein for their contributions to this project as well as my postgraduate journey over the past 3 years. I have learned so much from this department; the hardworking environment and sense of community that Prof. Amanda Krause has cultivated here are truly unmatched. A special thanks to Ms Monica Araujo for inspiring me to pursue this degree. Your belief in my potential and unwavering encouragement have been my guiding light throughout this journey. Without your support and inspiration, I would not have made it through this year. Thank you for being a constant source of strength and motivation.

To my friends and classmates, thank you for taking this journey with me. Thank you for supporting me through the tough times, celebrating with me through our triumphs and being the kind, beautiful souls that you are. I know you are all off to great places, as Dr Suess would say!

Lastly, thank you to my wonderful family for your unwavering support through my entire academic career. To my grandparents, Dr Bungar Naidoo and Mrs Kamala Naidoo, my aunts and uncle, Dr Thilomi Naidoo, Dr Somashree Moodley, Dr Serilla Moodley and Dr Kessendhra Naidoo, your steadfast support and sacrifices have provided me with the foundation to pursue this degree. To my siblings, Dylan and Reya, thank you for always bringing joy into my life, boosting my confidence, and reminding me of the importance of being a positive role model. I hope that you pursue your dreams with determination and confidence, knowing that you have the strength to achieve anything you set your minds to. And of course, to my parents, Dr Sevalen Moodley and Mrs Lovashini Moodley, thank you for instilling the importance of education in me. Your sacrifices and teachings have shaped the person I am today, and I carry your memory with me in every accomplishment. This journey is a tribute to the foundation you built for me, and I hope to honour your legacy through my work. I only wish you were both here to witness this.

TABLE OF CONTENTS

DECLARATION ..... **i**

CONTRIBUTION OF CANDIDATE ..... **ii**

DEDICATION ..... **iii**

PRESENTATIONS ARISING FROM THIS RESEARCH ..... **iv**

Abstract ..... **v**

ACKNOWLEDGEMENTS ..... **vi**

TABLE OF CONTENTS ..... **vii**

List of figures ..... **x**

List of tables ..... **xi**

List of abbreviations ..... **xii**

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

Abstract for journal submission **1**

1 INTRODUCTION **2**

2 METHODS AND MATERIALS **4**

2.1. Study Participants 4

2.2. Data generation using whole exome sequencing 4

2.3. Single nucleotide variant analysis: 4

2.3.1. Variant annotation 4

2.3.2. Variant prioritisation 4

2.3.3. Variant visualisation 4

2.3.4. Variant interpretation and classification 5

3 RESULTS 5

3.1. Patient demographics and WES data summary 5

3.2. Five variants met the filtering criteria 7

3.2.1. Identification of a likely pathogenic variant and VUS in the *OCA2* and *TYRP1* genes 9

3.2.2. Benign variants identified in the *OCA2, TYRP1* and *MCR1* genes 9

3.2.3. Common variants in the *OCA2* and *SLC24A5* genes identified in a majority of the patients 10

3.3. Summation of results 11

4 DISCUSSION 11

4.1. Variants identified 11

4.1.1. One pathogenic variant identified in a 2.7kb deletion negative patient 11

4.1.2. Classification of variant of uncertain significance identified in patient 16 12

4.1.3. Benign variants identified in patients 10, 11 and 13 13

4.1.4. Common variants and likely normal variation 13

4.2. Implications for clinical and genetic counselling 14

4.3. Limitations 14

4.4. Future directions 15

5 CONCLUSION 15

DATA AVAILABILITY STATEMENT 16

ETHICS STATEMENT 16

AUTHOR CONTRIBUTIONS 16

FUNDING 16

ACKNOWLEDGMENTS 16

CONFLICT OF INTEREST **16**

References **17**

Appendices **20**

Appendix A: Supplementary material 21

Appendix B: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M240749 29

Appendix C: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M210203 31

Appendix D: Approved research protocol 33

Appendix E: Plagiarism Declaration 51

Appendix F: Turnitin Report 52

Appendix G: Frontiers in Genetics Author guidelines 60

List of figures

Figure 1: Flow diagram illustrating variant analysis steps performed in this study. 4
Figure 2: MultiQC plots summarizing WES data for the eight patients..... 6

Supplementary figures

Supplementary figure 1: IGV visualisation of the *OCA2*: NM_000275.3: c.1239+2T>A variant in patient 11. 21

Supplementary figure 2: IGV visualisation of the *TYRP1*: NM_000550.3: c. 291G>C variant in patient 16.. 22

Supplementary figure 3: IGV visualisation of the *OCA2*: NM_000275.3: c.1007C>T variant in patient 10 and 13 22

Supplementary figure 4: IGV visualisation of the *TYRP1*: NM_000550.3: c.1557T>G variant in patient 13. 23

Supplementary figure 5: IGV visualisation of the *MC1R*: NM_002386.4: c.586T>C variant in patient 11. 23

List of tables

Table 1: Clinical information for patients as was available in the clinical files 5

Table 2: Variant shortlist from the cohort of African patients with *OCA2* 8

Supplementary tables

Supplementary table 1: Coverage of identified variants in each patient 21

Supplementary table 2: SpliceAI scores and their functional implications for the *OCA2*: c.1239+2T>A variant 24

List of abbreviations

Abbreviations	Definitions
aCGH	Array comparative genomic hybridisation
ACMG-AMP	American College of Medical Genetics and Genomics and the Association for Molecular Pathology
BAM	Binary alignment map
CNV	Copy number variant
GRCh37	Genome Reference Consortium Human Build 37
HGVS	Human Genome Variation Society
IGV	Integrative genomics viewer
LRMDA (C10ORF11)	Leucine-rich melanocyte differentiation associated gene
LRS	Long read sequencing
MAF	Minor allele frequency
MC1R	Melanocortin 1 receptor
MLPA	Multiplex ligation-dependent probe amplification
MRC-SIR	Medical Research Council - Self-Initiated Research award
NHLS	National health laboratory service
NGS	Next generation sequencing
OCA	Oculocutaneous albinism
OCA2	Oculocutaneous albinism type 2
OCA2	Oculocutaneous albinism type 2 gene
OCA3	Oculocutaneous albinism type 3
OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
ROCA	Rufous oculocutaneous albinism
SLC24A5	Solute carrier family 24 member 5
SLC45A2	Solute carrier family 45 member 2
SNV	Single nucleotide variant
TYR	Tyrosinase
TYRP1	Tyrosinase related protein 1
VCF	Variant call format
VEP	Variant effect predictor
VUS	Variant of uncertain significance
WES	Whole exome sequencing
WGS	Whole genome sequencing

Research report in the format of a "submissible" paper

To be submitted to the Frontiers in Genetics journal

Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant

Priya Moodley[1], Robyn A. Kerr [1], Phelelani T. Mpangase [2] and Thandiswa G. Ngcungcu [1]

[1] Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, The University of the Witwatersrand, Johannesburg, 2193, South Africa

[2] Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, The University of the Witwatersrand, Johannesburg, 2193, South Africa

Corresponding Authors:

Priya Moodley

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

[T]: 072 999 1074 [E]: priyamoodley71@gmail.com / 2111913@students.wits.ac.za

Thandiswa Ngcungcu

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

[T]: 073 847 8815 [E]: Thandiswa.Ngcungcu@wits.ac.za

Abstract for journal submission

Introduction: Oculocutaneous albinism (OCA) is a group of rare, inherited pigmentation disorders characterised by hypopigmentation of the skin and impaired vision. The most common form of OCA is oculocutaneous albinism type 2 (OCA2) which has a high prevalence in African populations. The most common cause of OCA2 in African individuals is a 2.7kb intragenic deletion variant in the *OCA2* gene. This deletion accounts for 78% of all pathogenic variants in the Southern African population, leaving 22% of alleles with an unknown (non-2.7kb deletion) variant. In order to investigate these 'other' alleles, this study aimed to identify pathogenic variants in African individuals with OCA2 that were known to be negative, or heterozygous, for the common *OCA2* 2.7 kb deletion.

Methods: Whole exome sequencing (WES) data was available for eight individuals who were suspected to have / clinically diagnosed with OCA2 and sent to the Molecular Genetics Diagnostic Laboratory, Division of Human Genetics, National Health Laboratory Service (NHLS) for genetic testing; individuals who tested negative or heterozygous for the 2.7kb deletion were included in this study. This WES data was analysed in an attempt to identify other possible *OCA2* pathogenic variants. This was done through bioinformatics analysis of WES data including variant annotation, prioritisation, visualisation and interpretation. Variants were annotated using Ensembl's Variant Effect Predictor (VEP) and then prioritised based on genes of interest, functionality, and minor allele frequency (MAF). The genes of interest included: *OCA2, TYR, MC1R, TYRP1, SLC45A2, SLC24A5,* and *C10ORF11/LRMDA*. Frameshift, missense, non-synonymous and splice site variants with a low MAF (<5%) were prioritised. Variants were visualised using the Integrative Genomics Viewer. Prioritised variants were classified according to the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guidelines.

Results: In a cohort of eight patients, this study identified one likely pathogenic variant, NM_000275.3: c.1239+2T>A, in the *OCA2* gene, in one patient and one variant of uncertain significance, NM_000550.3: c. 291G>C, in the *TYRP1* gene in another patient.

Conclusion: These results suggest that WES is a useful approach for identifying variants in the *OCA2* gene. It also suggests that clinical phenotyping can be inaccurate, and patients sent in for *OCA2* testing may have a different type of albinism - variants in the *TYRP1* gene are associated with OCA type 3. These findings underscore the need for further genetic investigation in larger cohorts to fully characterise the spectrum of OCA-associated variants in individuals negative or heterozygous for the 2.7kb deletion. Such insights are crucial for improving diagnostic accuracy and understanding the genetic basis of OCA in African populations.

Keywords: Oculocutaneous albinism type 2, OCA2, Southern African, whole exome sequencing, variant analysis

1 INTRODUCTION

Albinism is a group of rare, inherited pigmentation disorders characterised by hypopigmentation and impaired vision (Kromberg and Manga, 2018). Albinism exhibits an autosomal recessive pattern of inheritance. There are two main types of albinism, ocular albinism and oculocutaneous albinism (OCA). OCA is characterised by a lack of melanin pigment (hypopigmentation) from the skin, hair, and eyes from birth (King et al., 1995). There are eight types of OCA, each caused by variants in different genes (Kromberg et al., 2012). The most prevalent form of OCA worldwide is OCA2, due to its high frequency in individuals of African descent. In the Southern African population, the disorder has a frequency of 1/3900 (Kromberg and Jenkins, 1982). OCA results from defects in the melanin biosynthesis pathway, where variants in the *OCA2* gene can cause disruptions in the biochemical reactions that produce melanin, the pigment responsible for skin, hair, and eye coloration (Riley, 1997).

Individuals with OCA2 present with creamy white skin, impaired vision, nystagmus , strabismus and photophobia (Witkop, 1989). Some affected individuals present with skin patches called ephelides (large freckles) and hair that appears straw-coloured (yellow – blonde) or even light brown as some pigment may accumulate over time (Kromberg et al., 1989). Individuals with OCA2 lack melanin in the skin, leading to an increased risk of skin damage and skin cancers, commonly squamous cell carcinoma (Kromberg et al., 1989).

Oculocutaneous albinism type 2 results from variants in the *OCA2* gene on chromosome 15q (OMIM 203200). In 1994, Durham-Pierre and colleagues identified the 2.7kb deletion in the *OCA2* gene, which removes a single exon (exon 7) of the gene and leads to a frameshift deletion, resulting in a truncated and non-functional protein (Durham-Pierre et al., 1994). This 2.7kb deletion was found to be a founder variant in the Southern African population and accounts for 78% of all *OCA2* pathogenic variants in this population (Stevens et al., 1995). This means that 22% of pathogenic variants are not accounted for by the 2.7kb deletion, resulting in a proportion of individuals with OCA2 being either heterozygous for the deletion, or homozygous negative. The carrier frequency for this specific variant was found to be 1/40 in the Southern African population (Stevens et al., 1997).

Epistasis is the phenomenon of gene-gene interactions, whereby variants at one locus can influence the expression of variants present at another locus. This circumstance is usually observed between genes that function within the same pathway. The melanin biosynthesis pathway involves multiple genes such as *TYR, TYRP1, SLC24A5, MC1R* and *OCA2* (Chang et al., 2024). Changes to any of the genes in this pathway could affect melanin production and result in disease. For example, a previous study (Manga et al., 1997), found that the classic OCA2 phenotype was modified by a variant present at the *TYRP1* locus. Investigating the genes in this pathway may prove beneficial for researching non-2.7kb deletion variants.

Genetic testing for OCA2 can be performed using various molecular techniques, including targeted variant analysis, deletion and duplication analysis, exon mutation scanning, exon sequence analysis and sequence analysis of the entire coding region through techniques like next-generation sequencing (NGS). Internationally, comprehensive genetic testing for OCA2 often includes NGS-based methods that can analyze the entire gene and detect both single nucleotide variants (SNVs) and structural changes. These methods are widely available in high-resource settings, enabling accurate identification of causative variants. In South Africa, the only genetic test for OCA2 currently available through the NHLS is the gap-polymerase chain reaction (PCR) test developed by Durham-Pierre et. al., (1996) for the common 2.7kb deletion, which has been used since 1996 due to its efficacy and affordability (Kerr et al., 2000). Follow-up testing is not available due to the resource-constrained setting in South Africa.

Whole exome sequencing is a technique used to sequence the protein coding regions of the genome. This technique is useful because most known disease-causing variants are found in the coding regions, thus WES allows for detection of variants in all genes simultaneously at relatively low cost (Meienberg et al., 2016). The *OCA2* gene is large (about 345 kb) and consists of 24 exons. Sequencing all the exons of the *OCA2* gene, as well as the other OCA related genes, requires an efficient, cost-effective and high throughput technique such as WES (Hawkes et al., 2013). WES can detect SNVs and small insertions or deletions with high accuracy (Meienberg et al., 2016). This approach is useful for identifying genetic variants in individuals of African descent with OCA2 who are negative or heterozygous for the common African *OCA2* 2.7 kb deletion, as these individuals may carry other variants in *OCA2* or variants in other genes entirely, even though they are presenting with a similar phenotype (Meienberg et al., 2016).

Exploring the underlying variants that cause OCA2 will enable researchers and healthcare providers to provide better diagnosis and genetic counselling to individuals affected by this condition. Continued genetic studies should also aim to increase awareness of the biological nature of the disorder, serving to dispel myth and superstition that surrounds the aetiology of albinism in Africa.

This study is a continuation of a previous study (Mmbi et al., 2024), conducted on different patients, both of which fall under a larger parent study that is investigating the genetic basis of the different types of OCA in African populations. Mmbi and colleagues investigated individuals heterozygous for the 2.7kb deletion and identified one likely pathogenic variant, NM 000275.3: c.1503+5G>A, in the *OCA2* gene and two common rufous oculocutaneous albinism (ROCA) causing variants, NM 000550.3: c.497C>G and NM 000550.3: c.1103del in the *TYRP1* gene. In this present study, the aim was to increase the sample size and to further include individuals who were homozygous negative for the 2.7kb deletion, in an effort to contribute towards the knowledge of the variant spectrum underlying the genetic aetiology of OCA2.

2 METHODS AND MATERIALS

2.1. Study Participants

The study participants consisted of eight anonymised individuals of African descent with OCA2 who were found to be either negative (n=3) or heterozygous (n=5) for the 2.7kb deletion in the *OCA2* gene (confirmed by gap-PCR as part routine diagnostic testing).

2.2. Data generation using whole exome sequencing

The WES data for each of the participants had already been generated under the parent study (larger OCA2 study) and was available for analysis. The data was generated using the Ion Torrent™ Ion S5 and was available in the form of variant call format (VCF) files and binary map alignment (BAM) files. All data was aligned to the GRCh37 human reference genome build.

2.3. Single nucleotide variant analysis:

The WES data generated from the Ion Torrent™ S5 sequencer was analysed through various bioinformatics and clinical approaches as summarised by Figure 1:

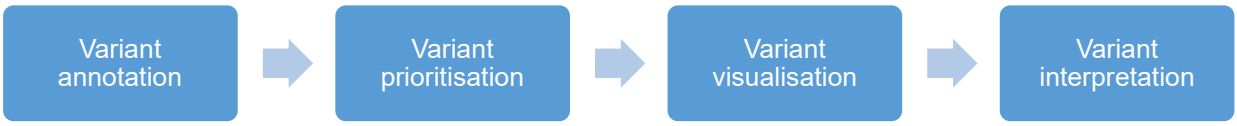


Figure 1: *Flow diagram illustrating variant analysis steps performed in this study.*

2.3.1. Variant annotation

Variants for each patient VCF file were annotated using Ensembl's VEP tool (release 112). The variant annotation assigned functional information and significance to variants in DNA, such as the type of variant and the level of pathogenicity of the variant (McCarthy et al., 2014). Information such as gene symbols, accession numbers, population frequency data and *in-silico* consequence prediction scores were also provided.

2.3.2. Variant prioritisation

Variants were then filtered and prioritised based on the gene of interest, functionality and allele frequency. The genes of interest include: *OCA2*, *TYR, MC1R, TYRP1, SLC45A2, SLC24A5, and C10ORF11/LRMDA* and variants in these genes were prioritised. Frameshift, missense, non-synonymous and splice site variants with a low MAF (<5%) were then prioritised.

2.3.3. Variant visualisation

A quality assessment and visualisation of the identified variants was performed using the Integrative Genomics Viewer (IGV) tool (version 2.17.4), which is used to visualise large genomic datasets (Thorvaldsdóttir et al., 2013). This tool was used to assess the depth of coverage and

confirm zygosity for each prioritised variant. A control sample was included and was a random, anonymised sample available from the Diagnostics Laboratory with no known diagnosis of OCA2.

2.3.4. Variant interpretation and classification

Prioritised variants were classified based on the ACMG-AMP guidelines for variant interpretation (Supplementary document 1). The ACMG-AMP guidelines provide a framework for classifying genetic variants into different categories based on evidence such as population data, computational predictions, functional studies, segregation with disease, *de novo* occurrence, and allelic data. After combining the evidence and assessing their strength, the prioritised variants were classified into five categories: pathogenic, likely pathogenic, variant of uncertain significance (VUS), likely benign, and benign (Richards et al., 2015).

3 RESULTS

3.1. Patient demographics and WES data summary

Patient clinical and demographic information for the eight individuals was available from their test request forms sent to the Diagnostic Laboratory at the time of the test request (Table 1).

Table 1: Clinical information for patients as was available in the clinical files

Patient ID	Ethnicity	Phenotype	2.7kb deletion	Clinical information*
P009	Black	Affected	Heterozygous	Generalised hypopigmentation, slight grey hair, white eyelashes/eyebrows
P010	Black	Affected	Heterozygous	Ephelides
P011	Black	Affected	Negative	No data provided
P013	Black	Affected	Heterozygous	Affected with OCA2
P014	Black	Affected	Heterozygous	Patient is said to present with OCA, one unidentified variant
P015	Black	Affected	Negative	Generalised hypopigmentation, slight grey hair, white eyelashes/eyebrows
P016	Black	Affected	Heterozygous	OCA2 phenotype
P017	Black	Affected	Negative	No data provided

*Clinical information available on request form provided by clinician

The previously generated WES produced data with a mean depth coverage of 4.8x – 5.5x and a mean read length of 190 base pairs. For all samples, at least 99% of the sequenced reads were aligned to the GRCh37 reference sequence. The VEP annotation of each patient file produced between 41000 to 62000 variants per patient. The most common variants per patient were SNVs, specifically intronic variants, synonymous variants and missense variants, respectively (Figure 2).

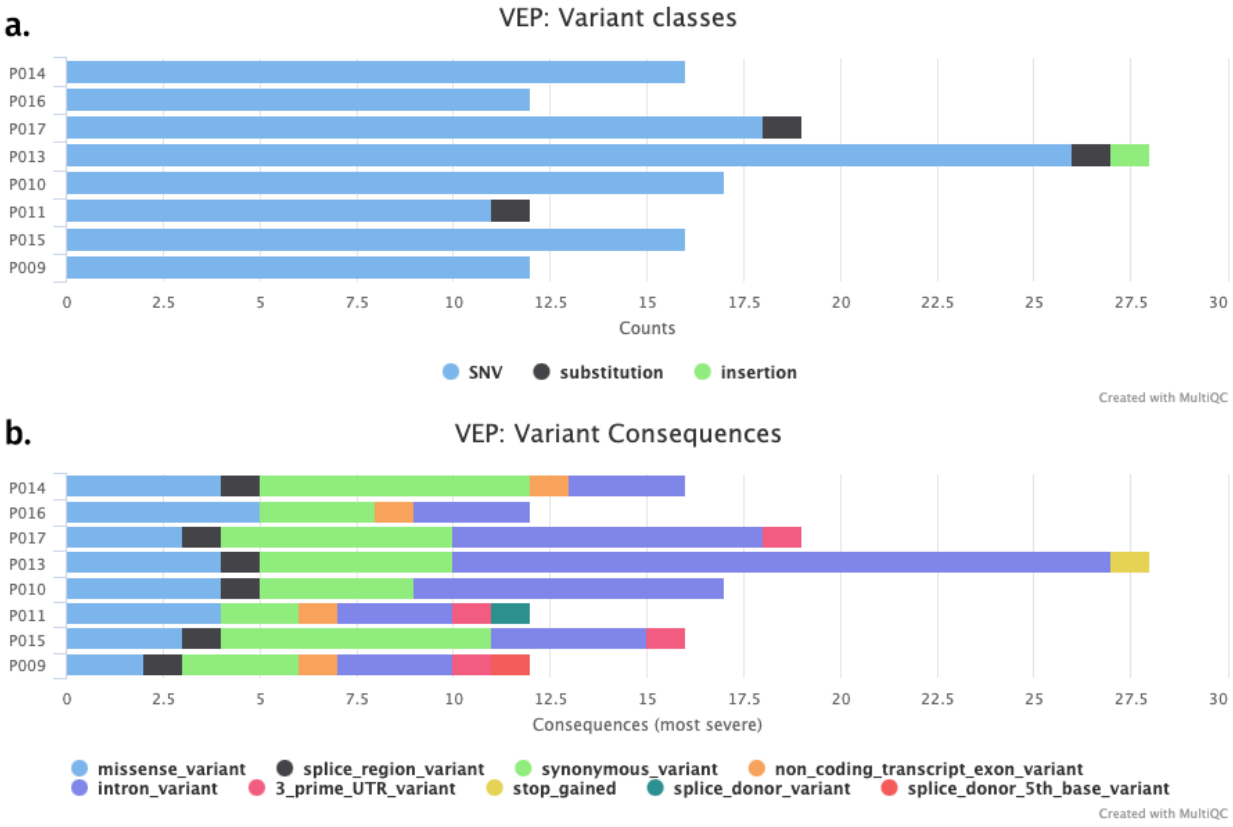


Figure 2: *MultiQC plots summarizing WES data for the eight patients.*

a) Distribution of the variant classes identified through the VEP annotation across all eight patients.

b) Type of variant consequences called from the VEP annotation for each patient in the South African cohort.

After the annotation, variants were filtered based on the genes of interest, functionality and impact and MAF (<5%). Only variants in the *OCA2*, *TYR, MC1R, TYRP1, SLC45A2, SLC24A5, and C10ORF11/LRMDA* genes were prioritised for further analysis. This filtering step resulted in 162 prioritised variants. The next filtering step (functionality and impact) produced 11 variants for further assessment: nine missense variants, one stop-gain and one splice donor variant. Lastly, only variants with a low MAF of less than 5% were prioritised, resulting in a shortlist of five prioritised variants (Table 2).

The five variants that met the filtering criteria were identified in the *MC1R, OCA2* and *TYRP1* genes. These five variants were visualised using IGV and were not found in homopolymer regions nor located in regions with many insertions and deletions. The coverage of these regions were all greater than 44x (Supplementary table 1).

3.2. Five variants met the filtering criteria

No variants that met the filtering criteria were identified for patients 9, 14, 15 and 17. The variant analysis of the eight South African patients resulted in a shortlist of five variants across four patients (Table 1). A likely pathogenic variant in the was identified in the *OCA2* gene (NM_000275.3: c.1239+2T>A) in patient 11 and a VUS was identified in patient 16, in the *TYRP1* gene (NM_000550.3: c.291G>C). Three variants were identified and classified as benign in the *MC1R, OCA2* and *TYRP1* genes.

Table 2: Variant shortlist from the cohort of African patients with OCA2

Patient	Gene	Location	HGVSc	HGVSp	Minor allele frequency*	Variant consequence	Zygosity	ACMG codes	Classification
11	*MC1R*	16:89986252	NM_002386.4: c.586T>C	p.Phe196Leu	0.0339	Missense	Heterozygous	BS1, BS2, BP6	Benign
10 ,13	*OCA2*	15:28259959	NM_000275.3: c.1007C>T	p.Ala336Val	0.0246	Missense	Heterozygous	BS1, BS2, BP4, BP6	Benign
11	*OCA2*	15:28231731	NM_000275.3: c.1239+2T>A	N/A	Not observed	Splice donor	Heterozygous	PVS1, PM2, PP3, PP5	Likely Pathogenic
13	*TYRP1*	9:12709125	NM_000550.3: c.1557T>G	p.Tyr519Ter	0.0386	Stop gain	Heterozygous	BS1, BS2, BP6, PVS1	Benign
16	*TYRP1*	9:12694287	NM_000550.3: c.291G>C	p.Arg97Ser	0.0000	Missense	Heterozygous	PM2	VUS

*Minor allele frequency in normal African population (the African/African American ancestry group as defined in gnomAD v2.1.1))

Version 2.1.1 of gnomAD was used for this analysis because the WES data generated for this project was aligned to the reference genome GRCh37. This version of gnomAD is the only one compatible with GRCh37 for SNV analysis, ensuring accurate and consistent allele frequency estimations relevant to this project.

This table summarises the SNVs identified in three OCA-related genes (*MC1R, OCA2*, and *TYRP1*) across six different patients, highlighting their potential clinical significance. Each variant is described with its coding DNA (HGVSc) and protein (HGVSp) changes using standardised Human Genome Variation Society (HGVS) nomenclature. Minor allele frequencies within the African population are provided, illustrating the prevalence of these variants in this population. The table also outlines the functional consequences of each variant, such as missense variants in *MC1R*, *OCA2* and *TYRP1*, a splice donor alteration in *OCA2*, and a stop-gain variant in *TYRP1*. Each variant has been classified according to ACMG-AMP guidelines, with specific codes indicating the evidence level supporting each classification. Variants are classified as likely pathogenic, a VUS or benign.

3.2.1. Identification of a likely pathogenic variant and VUS in the *OCA2* and *TYRP1* genes

3.2.1.1. A likely pathogenic variant identified in the *OCA2* in patient 11

A splice donor variant, in the *OCA2* gene*,* NM_000275.3: c.1239+2T>A, was identified in patient 11. The variant results in abnormal splicing of the transcript. This type of variant can disrupt normal RNA splicing, often resulting in an abnormal protein due to exon skipping, intron retention, or a frameshift, potentially leading to a non-functional or degraded protein. The variant was visualised using IGV, confirming its heterozygous zygosity (Supplementary figure 1). Code PVS1 can be applied to this variant because a splice donor variant is a null variant, and it occurs in the *OCA2* gene where loss of function is a known mechanism of disease. This variant has not been reported in population databases such as gnomAD, thus the ACMG-AMP code PM2 was applied to this variant (Table 2). Several *in silico* prediction tools on Varsome (27/29) predict the variant to be pathogenic, thus PP3 was applied. The code PP5 was also applied because this variant was reported as likely pathogenic on ClinVar, supported by previous publications that provide evidence for the pathogenicity of splice-site variants and OCA2. Specifically, these studies highlighted the role of splice-site disruptions in disease causation (Baralle and Baralle, 2005) and include comprehensive analyses of *OCA2* variants in affected individuals (Preising et al., 2011; Wei et al., 2010). While these publications do not explicitly analyse this specific splice donor variant, they support the general classification framework applied to *OCA2* splice-site variants. This variant was therefore classified as likely pathogenic based on this information.

3.2.1.2. A variant of unknown significance identified in the *TYRP1* gene

A missense variant, in the *TYRP1* gene*,* NM_000550.3: c. 291G>C (p.Arg97Ser), was identified in patient 16. This variant results in a change of amino acid at position 97. This type of variant alters the protein sequence and possibly its function. The variant was visualised using IGV, confirming its heterozygous zygosity (Supplementary figure 2). This variant has been reported in population databases, such as gnomAD, with an African frequency of 0.000, thus the ACMG-AMP codes PM2 was applied to this variant (Table 2). European populations were also reported to have a frequency of 0.000 on gnomAD. This code is applied when the allele frequency is extremely low in the population. No further information is available for this variant; therefore, it was classified as a VUS.

3.2.2. Benign variants identified in the *OCA2, TYRP1* and *MCR1* genes

3.2.2.1. Benign missense variant identified in the *OCA2* gene in two patients

A missense variant, in the *OCA2* gene*,* NM_000275.3: c.1007C>T (p.Ala336Val)*,* was identified in patient 10 and 13. This variant results in a change of amino acid at position 336. This type of variant may alter the protein sequence and thus its function. The variant was visualised using IGV, confirming its heterozygous zygosity (Supplementary figure 3). This variant has been reported in population databases such as gnomAD, with an African frequency of 0.02456, thus the ACMG-AMP codes BS1 and BS2 were applied to this variant (Table 2). These codes are applied when the

allele frequency is greater than expected and the variant was observed in a homozygous state in population databases more than expected for the disease (OCA2). Code BP4 was applied because multiple *in silico* prediction tools on Varsome (54/58) predict that the variant is benign. The code BP6 was also applied because this variant was reported as benign or likely benign on UniProt and ClinVar with a two-star rating (indicating multiple submitters agree with the classification) and supporting publications (Kamaraj and Purohit, 2014; Thomas et al., 1993). This variant was therefore classified as benign based on this information.

3.2.2.2. Benign *TYRP1* stop gain variant identified in patient 13

A stop gain variant, in the *TYRP1* gene, NM_000550.3: c.1557T>G (p.Tyr519*), was identified in patient 13. This variant results in a premature stop codon and thus early termination of the amino acid sequence, and truncation of the protein. This type of variant usually results in non-functional or degraded protein. Code PVS1 can be applied to this variant because a stop gain variant is a null variant, and it occurs in the *TYRP1* gene where loss of function is a known mechanism of disease. The variant was visualised using IGV, confirming its heterozygous zygosity (Supplementary figure 4). However, this variant has been reported in population databases such as gnomAD, with an African frequency of 0.03856, thus the ACMG-AMP codes BS1 and BS2 were applied to this variant (Table 2), strongly supporting a benign classification. These codes are applied when the allele frequency is common in the population and when the variant has been observed in a homozygous state in population databases more often than expected. The code BP6 was also applied because this variant was reported as benign by six independent submitters in ClinVar with a two-star rating. This variant was therefore classified as benign based on this information.

3.2.2.3. Benign missense variant identified in the *MC1R* gene in patient 11

This missense variant, in the *MC1R* gene, NM_002386.4: c.586T>C (p.Phe196Leu), was identified in patient 11. This variant results in a change of amino acid at position 196 (p.Phe196Leu). This type of variant may alter the protein sequence and thus its function. The variant was visualised using IGV, confirming its heterozygous zygosity (Supplementary figure 5). This variant has been reported in population databases such as gnomAD, with an African frequency of 0.03391, thus the ACMG-AMP codes BS1 and BS2 were applied to this variant (Table 2). The code BP6 was also applied because this variant was reported as benign or likely benign on ClinVar with a two-star rating and the support of several publications (Demenais et al., 2010; Puig-Butillé et al., 2013; Thomas et al., 1993). This variant was therefore classified as benign based on this information.

3.2.3. Common variants in the *OCA2* and *SLC24A5* genes identified in a majority of the patients

Analysis of this cohort also revealed two common missense variants observed across multiple individuals. These include one variant in *OCA2*, NM_000275.3: c.913C>T (present in 5 of 8

patients) and one variant in *SLC24A5*, NM_205850.3: c.331A>G (present in 7 of 8 patients). Based on their high MAF in the gnomAD database for African populations, these variants are likely benign polymorphisms or represent normal population variation. Specifically, the *OCA2*: c.913C>T variant has a reported MAF of 0.1385, while *SLC24A5*: c.331A>G has a MAF of 0.8088.

3.3. Summation of results

No variants meeting the filtering criteria were identified for patients 9, 14, 15, and 17. For patient 11, who is negative for the 2.7 kb deletion, a likely pathogenic splice donor variant (NM_000275.3: c.1239+2T>A) was identified. Additionally, a benign variant was found in the *MC1R* gene for this patient. In patient 16, who is heterozygous for the 2.7 kb deletion, a VUS in the *TYRP1* gene (NM_000550.3: c.291G>C) was identified. Patients 10 and 13 are both heterozygous for the 2.7 kb deletion, but no second variant (pathogenic or likely pathogenic) was identified in the *OCA2* gene or any other OCA-related genes; only benign variants were found in these patients.

4 DISCUSSION

This study aimed to identify germline variants in African patients clinically diagnosed with OCA2. Whole exome sequencing data produced using the Ion Torrent™ Ion S5 system, was available for analysis. Five variants in the *MC1R*, *OCA2* and *TYRP1* genes met the filtering criteria and were analysed further. The study was successful in identifying one likely pathogenic variant in *OCA2* in patient 11 and one VUS in *TYRP1* in patient 16.

4.1. Variants identified

4.1.1. One likely pathogenic variant identified in a 2.7kb deletion negative patient

A likely pathogenic variant was identified in patient 11 in the *OCA2* gene (*OCA2:* NM_000275.3: c.1239+2T>A). It is a splice-donor variant, which has been previously reported and classified as likely pathogenic on ClinVar with a one-star rating (Preising et al., 2011; Wei et al., 2010). Splicing variants in the *OCA2* gene can disrupt the normal processing of mRNA, leading to abnormal or incomplete protein production. This can cause exon skipping, intron retention, or the creation of a frameshift, resulting in a truncated or non-functional OCA2 protein (Baralle and Baralle, 2005). Since the OCA2 protein plays a key role in melanin production, these splicing disruptions can reduce or eliminate melanin synthesis, leading to the hypopigmentation and visual impairments characteristic of OCA2 (Blakes et al., 2022; Kromberg and Manga, 2018). Code PVS1 was applied to splice-donor variant which is considered a null variant. It occurs in the *OCA2* gene, where loss of function is a well-established disease mechanism. Further analysis, using SpliceAI, supports this by demonstrating an extremely high likelihood (99%) that the original donor splice site is disrupted and a high likelihood (91%) of a new donor splice site forming, consistent with significant splicing

disruption (Supplementary table 2). This prediction from SpliceAI along with several *in silico* prediction tools on Varsome (27/29) predict the variant to be pathogenic, so PP3 was applied. The code PM2 was applied because the variant is absent in population databases such as gnomAD. This splice-donor variant is classified as likely pathogenic in ClinVar, supported by references to prior studies that highlight the role of splicing disruptions in *OCA2* (Baralle and Baralle, 2005; Preising et al., 2011; Wei et al., 2010). Although these publications do not specifically examine this variant, they support the pathogenicity framework for splice-site variants in *OCA2*. ClinVar assigns this variant a one-star rating, which means that its classification is based on evidence submitted by a single submitter (Labcorp Genetics, formerly Invitae), that has yet to be reviewed by an independent expert. While this one-star rating highlights the need for further validation, it does not detract from its clinical relevance, given the strong evidence supporting the splice-site disruption in the *OCA2* gene. As PP5 only provides supporting evidence rather than pivotal justification, it contributes to the weight of evidence for pathogenicity (Richards et al., 2015). Unfortunately, clinical information was unavailable for this patient (P011), thus no codes associated with specific phenotype could be applied. For example, if this case is followed up and OCA2 specific clinical features can be confirmed, PP4 could also be applied to this variant providing more supporting evidence of pathogenicity. The inclusion of PVS1, PM2, PP3 and PP5 ACMG-AMP criteria in its classification reflect strong evidence that this variant could disrupt gene function (Riggs et al., 2020).

4.1.2. Classification of variant of uncertain significance identified in patient 16

One VUS, *TYRP1* c.291G>C (p.Arg97Ser), was identified, which has not been previously reported and has an African minor allele frequency of zero in gnomAD. Variants in the *TYRP1* gene are usually associated with OCA type 3 (Manga et al., 1997; Patel et al., 2021). This VUS could potentially play a role in the phenotype, although further functional studies would be necessary to determine its exact consequence and clinical relevance. Code PVS1 can be applied to this variant because a stop gain variant is a null variant, and it occurs in the *TYRP1* gene where loss of function is a known mechanism of disease. This patient may have been clinically misdiagnosed due to the significant overlap between different types of OCA and often only molecular testing can differentiate between the types (Kromberg and Manga, 2018). Similar findings were reported by Mmbi et al., (2024) who highlighted the challenges in accurately distinguishing OCA phenotypes based on clinical features alone, particularly in African populations where limited molecular data is available. It is possible that this patient has oculocutaneous albinism type 3 (OCA3) - although this is unlikely as a second variant was not found. Another possibility for this variant is that it has an epistatic effect on variation in the *OCA2* gene, or another OCA-related gene, and the combination of variants produces an OCA clinical picture. The absence of any information on this variant highlights the challenges associated with VUS classification, particularly in understudied populations where comprehensive variant data may be lacking.

4.1.3. Benign variants identified in patients 10, 11 and 13

Three out of five prioritised variants in this study were classified as benign (*MC1R*: NM_002386.4: c.586T>C, *OCA2*: NM_000275.3: c.1007C>T and *TYRP1*: NM_000550.3: c.1557T>G). Each of these benign variants had been previously classified as benign on ClinVar (with supporting publications) and appear in the population with higher frequency than expected, providing further evidence of their non-pathogenicity. The *MC1R* variant was only found in patient 11, the *OCA2* variant was found in patients 10 and 13 and the *TYRP1* variant was found in patient 13. The consistent benign classification of these variants across different databases and clinical reports underscores the utility of existing variant databases in streamlining clinical decision-making (Riggs et al., 2020).

The *TYRP1* gene has been implicated in OCA3 in the Southern African population, with specific variants leading to loss of function and distinct pigmentation phenotypes (Manga et al., 1997). There are two common variants described by Manga et al., (1997): a nonsense variant (NM_000550.3: c.497C>A) and a frameshift variant (NM_000550.3: c.1118delA). Both these variants result in premature stop codons which truncate the protein. In contrast, the identified stop-gain variant (NM_000550.3: c.1557T>G) is classified as benign based on allele frequency and lacks clear pathogenicity evidence. This variant has an African allele frequency of 0.03856 in gnomAD, which is significantly higher compared to its frequency in the European population (0.00001555). This suggests that the variant is most likely a population-specific polymorphism rather than a pathogenic variant in African individuals. However, while its high frequency in Africans supports its benign classification, its role as a potential modifier or its population-specific effects cannot be entirely excluded without further study. Given the high prevalence of *TYRP1* variants in hypopigmentation disorders, even benign variants could contribute to phenotypic variability or represent adaptation within specific genetic backgrounds.

4.1.4. Common variants and likely normal variation

After analysis of this South African cohort with query-OCA2, two common missense variants were observed across multiple individuals. These variants, located in *OCA2*: NM_000275.3: c.913C>T (in 5/8 patients) and *SLC24A5*: NM_205850.3: c.331A>G (in 7/8 patients), are most likely benign polymorphisms and represent normal population variation. This is supported by the African minor allele frequencies observed in the population database, gnomAD. The *OCA2*: c.913C>T variant, has an allele frequency of 0.1385 and the *SLC24A5*: c.331A>G variant, has an allele frequency of 0.8088. Both of these frequencies are relatively high and thus considered to be common since rare variants are described to have a MAF of less than 0.01 (1%) (Goswami et al., 2021). This highlights the importance of population-specific data, as variants that appear frequently within one

population might be erroneously classified as significant if diverse-population data is not available (Choudhury et al., 2014; Christen and Schnell, 2023; Gudmundsson et al., 2021).

4.2. Implications for clinical and genetic counselling

The identification of both benign and pathogenic variants provides valuable insights into the genetic background of OCA2 within Southern African populations. The commonality of the identified benign variants and their previous classifications help mitigate potential over-interpretation in clinical settings. However, the detection of a VUS and a likely pathogenic variant highlights the need for careful interpretation and potential expansion of molecular testing for OCA2. This expanded knowledge of *OCA2* variants allows for more accurate and supportive genetic counselling, which improves care and empowers patients and families with actionable information (Patch and Middleton, 2018). These findings underscore the significance of establishing well-curated and diverse population databases to avoid misclassification and ensure accurate variant interpretation.

4.3. Limitations

In this cohort, no significant variants were identified in six patients (only benign variants were found), and even in the two patients with the VUS or the likely pathogenic variant, an additional variant is required for clinical confirmation and diagnosis, as OCA2 is a recessive disorder. Whole exome sequencing would not be able to detect the second variant if it were in a regulatory region, or a deep intronic variant, or if the variation was epigenetic (Yao et al., 2017). Furthermore, WES only focuses on protein-coding regions, limiting its ability to detect variants in non-coding regions like regulatory elements and introns, despite their potential impact on gene function and disease (Meienberg et al., 2016). There are also other challenges with WES such as incomplete exon coverage and its limited ability to detect structural variants (copy number variants - CNVs). A further limitation of this study was the inability to perform CNV analysis due to technical and resource constraints, which prevented the detection of potential structural variants. A previous study (Mmbi, 2024) was also unable to detect CNVs (including the 2.7 kb deletion) using Ion Reporter software, suggesting that this software is not well designed/adapted for smaller CNVs. Additionally, since only a small subset of genes was analysed within the WES data, it is possible that variants in other important pigmentation-related genes, genes involved in epistasis, and modifier genes, were missed. Lastly, since all the patients included in this study were of African ancestry, there were challenges in variant interpretation. This is because some of the identified variants are poorly characterised in genetic databases and published literature, due to the lack of African-specific population data.

4.4. Future directions

While this study focused on a limited subset of seven genes, future research could expand to a larger virtual gene panel or whole genome sequencing (WGS) to capture variants in non-coding regions and other genes involved in pigmentation. This broader approach could uncover novel variants and regulatory elements that impact the OCA2 phenotype, providing a more comprehensive genetic view of OCA2 (Diallo et al., 2024). Additionally, pathway analyses of the identified genes could reveal genetic interactions and modifier effects, offering deeper insights into the genetic basis of OCA. For further investigation of the VUS in *TYRP1*, functional studies are recommended. Techniques such as gene editing or expression assays could help determine the impact this variant has on protein function or gene expression, adding valuable context to the genetic findings and aiding in future variant classification (Guidugli et al., 2013). The study highlighted the need for more population-specific genetic data. Developing a comprehensive and robust database of African-specific allele frequencies would reduce uncertainty in variant interpretation and improve classification accuracy for African patients. This could prevent the misclassification of common polymorphisms as pathogenic variants (Christen and Schnell, 2023). Given the limitations of WES and Ion Reporter for detecting CNVs, future studies could incorporate additional CNV-specific techniques, such as quantitative PCR (qPCR), multiplex ligation-dependent probe amplification (MLPA), or array comparative genomic hybridization (aCGH), to validate and confirm any CNV findings (Morice-Picard et al., 2014). This could improve CNV detection accuracy, especially for known deletions like the 2.7 kb deletion in *OCA2*, which could not be confirmed in related studies using Ion Reporter (Mmbi et al., 2024). Additionally, another option for CNV analysis could be long read sequencing (LRS) if there is an option to curate new sequencing data from the original samples. LRS is ideal for CNV analysis because it provides continuous, extensive reads that accurately detect and characterise complex and large structural variants, even in repetitive regions (Yuan and Jia, 2024).

5 CONCLUSION

This study successfully identified a likely pathogenic variant in *OCA2* and a VUS in *TYRP1*, offering further insights into the genetic landscape of OCA2 in a South African cohort. Findings are in-keeping with the knowledge that the 2.7kb deletion is the common *OCA2* variant in this population, and besides this founder variant, few other pathogenic variants can be found. Common benign variants were also detected, highlighting the importance of population-specific data to avoid misclassifications. Limitations of WES, including its inability to detect non-coding variants and CNVs, suggest the use of broader techniques like aCGH or WGS in future studies. Expanding research to include additional genes, as well as analysis of the melanin biosynthesis pathway, could provide a more comprehensive understanding of OCA2 and the genetics of hypopigmentation, while population-specific databases, enriched with African specific data, would

enhance the accuracy of variant interpretation in African populations. Furthermore, this study contributes to scientific innovation in Southern Africa by advancing research in genomic medicine and encouraging the development of resources that address unique genetic disorders within this under-represented region, which may ultimately improve diagnostic precision and health outcomes.

DATA AVAILABILITY STATEMENT

The datasets utilised in this study are not available to the public due to ethical and privacy restrictions.

ETHICS STATEMENT

The study was approved by the Human Research Ethics Committee of the University of the Witwatersrand (M240749).

AUTHOR CONTRIBUTIONS

Dr. Thandiswa Ngcungcu conceptualised and designed the study. Bioinformatics analysis was performed by Dr. Phelelani T. Mpangase and Priya Moodley. Priya Moodley prepared the initial draft of the manuscript. Dr. Robyn Kerr and Dr. Thandiswa Ngcungcu provided theoretical guidance. All authors — Dr. Thandiswa Ngcungcu, Dr. Robyn Kerr, Dr. Phelelani T. Mpangase and Priya Moodley — critically revised the manuscript and approved the final version for submission.

FUNDING

This study was funded through the MRC SIR Grant (2021) awarded to Dr Thandiswa Ngcungcu for the parent project titled “The identification of causal mutations of oculocutaneous albinism type 2 (OCA2) in individuals of African descent”.

ACKNOWLEDGMENTS

We thank the Division of Human Genetics, National Health Laboratory Service (NHLS) for providing samples used for this study.

CONFLICT OF INTEREST

The authors confirm that the study was conducted without any conflicts of interest.

References

Baralle, D., Baralle, M., 2005. Splicing in action: assessing disease causing sequence changes. Journal of Medical Genetics 42, 737. https://doi.org/10.1136/jmg.2004.029538

Blakes, A.J.M., Wai, H.A., Davies, I., Moledina, H.E., Ruiz, A., Thomas, T., et al., 2022. A systematic analysis of splicing variants identifies new diagnoses in the 100,000 Genomes Project. Genome Medicine 14, 79. https://doi.org/10.1186/s13073-022-01087-x

Chang, Y., Wu, S., Li, J., Bao, H., Wu, C., 2024. Identification of Candidate Genes for Red-Eyed (Albinism) Domestic Guppies Using Genomic and Transcriptomic Analyses. International Journal of Molecular Sciences 25, 2175. https://doi.org/10.3390/ijms25042175

Choudhury, A., Hazelhurst, S., Meintjes, A., Achinike-Oduaran, O., Aron, S., Gamieldien, J., et al., 2014. Population-specific common SNPs reflect demographic histories and highlight regions of genomic plasticity with functional relevance. BMC Genomics 15, 437. https://doi.org/10.1186/1471-2164-15-437

Christen, P., Schnell, R., 2023. Thirty-three myths and misconceptions about population data: from data capture and processing to linkage. International Journal of Population Data Science 8. https://doi.org/10.23889/ijpds.v8i1.2115

Demenais, F., Mohamdi, H., Chaudru, V., Goldstein, A.M., Newton Bishop, J.A., Bishop, D.T., et al., 2010. Association of MC1R variants and host phenotypes with melanoma risk in CDKN2A mutation carriers: a GenoMEL study. J Natl Cancer Inst 102, 1568–1583. https://doi.org/10.1093/jnci/djq363

Diallo, M., Sylla, O., Sidibé, M.K., Plaisant, C., Mercier, E., Sequeira, A., et al., 2024. Genotypic spectrum of albinism in Mali. Pigment Cell & Melanoma Research 37, 752–761. https://doi.org/10.1111/pcmr.13175

Durham-Pierre, D., Gardner, J.M., Nakatsu, Y., King, R.A., Francke, U., Ching, A., et al., 1994. African origin of an intragenic deletion of the human P gene in tyrosinase positive oculocutaneous albinism. Nat Genet 7, 176–179. https://doi.org/10.1038/ng0694-176

Goswami, C., Chattopadhyay, A., Chuang, E.Y., 2021. Rare variants: data types and analysis strategies. Annals of Translational Medicine 9, 961. https://doi.org/10.21037/atm-21-1635

Gudmundsson, S., Singer-Berk, M., Watts, N.A., Phu, W., Goodrich, J.K., Solomonson, M., et al., 2021. Variant interpretation using population databases: Lessons from gnomAD. Human Mutation 43, 1012. https://doi.org/10.1002/humu.24309

Guidugli, L., Carreira, A., Caputo, S.M., Ehlen, A., Galli, A., Monteiro, A.N., et al., 2013. Functional assays for analysis of variants of uncertain significance in BRCA2. Human mutation 35, 151. https://doi.org/10.1002/humu.22478

Hawkes, J.E., Cassidy, P.B., Manga, P., Boissy, R.E., Goldgar, D., Cannon-Albright, L., et al., 2013. Report of a Novel OCA2 Gene Mutation and an Investigation of Two OCA2 Variants on Melanoma Predisposition in a Familial Melanoma Pedigree. J Dermatol Sci 69, 30–37. https://doi.org/10.1016/j.jdermsci.2012.09.016

Kamaraj, B., Purohit, R., 2014. Computational screening of disease-associated mutations in OCA2 gene. Cell Biochem Biophys 68, 97–109. https://doi.org/10.1007/s12013-013-9697-2

Kerr, R., Stevens, G., Manga, P., Salm, S., John, P., Haw, T., et al., 2000. Identification of P gene mutations in individuals with oculocutaneous albinism in sub-Saharan Africa. Hum Mutat 15, 166–172. https://doi.org/10.1002/(SICI)1098-1004(200002)15:2<166::AID-HUMU5>3.0.CO;2-Z

King, R.A., Hearing, V.J., Creel, D.J., Oetting, W.S., 1995. Albinism, in: Valle, D.L., Antonarakis, S., Ballabio, A., Beaudet, A.L., Mitchell, G.A. (Eds.), The Metabolic and Molecular Bases of Inherited Disease. McGraw-Hill Education, New York, NY.

Kromberg, J., Manga, P., 2018. Albinism in Africa. Academic Press, London, United Kingdom.

Kromberg, J.G., Jenkins, T., 1982. Prevalence of albinism in the South African negro. S Afr Med J 61, 383–386.

Kromberg, J.G.R., Bothwell, J., Kidson, S.H., Manga, P., Kerr, R., Jenkins, T., 2012. TYPES OF ALBINISM IN THE BLACK SOUTHERN AFRICA POPULATION. East Afr Med J 89, 20–27.

Kromberg, J.G.R., Castle, D., Zwane, E.M., Jenkins, T., 1989. Albinism and skin cancer in Southern Africa. Clinical Genetics 36, 43–52. https://doi.org/10.1111/j.1399-0004.1989.tb03365.x

Manga, P., Kromberg, J.G., Box, N.F., Sturm, R.A., Jenkins, T., Ramsay, M., 1997. Rufous oculocutaneous albinism in southern African Blacks is caused by mutations in the TYRP1 gene. Am J Hum Genet 61, 1095–1101. https://doi.org/10.1086/301603

McCarthy, D.J., Humburg, P., Kanapin, A., Rivas, M.A., Gaulton, K., Cazier, J.-B., et al., 2014. Choice of transcripts and software has a large effect on variant annotation. Genome Med 6, 26. https://doi.org/10.1186/gm543

Meienberg, J., Bruggmann, R., Oexle, K., Matyas, G., 2016. Clinical sequencing: is WGS the better WES? Hum Genet 135, 359–362. https://doi.org/10.1007/s00439-015-1631-9

Mmbi, P., Kerr, R., Mpangase, P., Ngcungcu, T. 2024. Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of African descent through exome sequencing. Master's research report. Johannesburg. University of the Witwatersrand.

Morice-Picard, F., Lasseaux, E., Cailley, D., Gros, A., Toutain, J., Plaisant, C., et al., 2014. High-resolution array-CGH in patients with oculocutaneous albinism identifies new deletions of the TYR, OCA2, and SLC45A2 genes and a complex rearrangement of the OCA2 gene. Pigment Cell Melanoma Res 27, 59–71. https://doi.org/10.1111/pcmr.12173

Patch, C., Middleton, A., 2018. Genetic counselling in the era of genomic medicine. British Medical Bulletin 126, 27. https://doi.org/10.1093/bmb/ldy008

Patel, M.H., Dolinska, M.B., Sergeev, Y.V., 2021. Tyrp1 Mutant Variants Associated with OCA3: Computational Characterization of Protein Stability and Ligand Binding. International Journal of Molecular Sciences 22, 10203. https://doi.org/10.3390/ijms221910203

Preising, M.N., Forster, H., Gonser, M., Lorenz, B., 2011. Screening of TYR, OCA2, GPR143, and MC1R in patients with congenital nystagmus, macular hypoplasia, and fundus hypopigmentation indicating albinism. Molecular Vision 17, 939.

Puig-Butillé, J.A., Carrera, C., Kumar, R., Garcia-Casado, Z., Badenas, C., Aguilera, P., et al., 2013. Distribution of MC1R variants among melanoma subtypes: p.R163Q is associated with lentigo maligna melanoma in a Mediterranean population. Br J Dermatol 169, 804–811. https://doi.org/10.1111/bjd.12418

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al., 2015. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 17, 405–424. https://doi.org/10.1038/gim.2015.30

Riggs, E.R., Andersen, E.F., Cherry, A.M., Kantarci, S., Kearney, H., Patel, A., et al., 2020. Technical standards for the interpretation and reporting of constitutional copy number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). Genet Med 22, 245–257. https://doi.org/10.1038/s41436-019-0686-8

Riley, P.A., 1997. Melanin. The International Journal of Biochemistry & Cell Biology 29, 1235–1239. https://doi.org/10.1016/S1357-2725(97)00013-7

Stevens, G., Ramsay, M., Jenkins, T., 1997. Oculocutaneous albinism (OCA2) in sub-Saharan Africa: distribution of the common 2.7-kb P gene deletion mutation. Hum Genet 99, 523–527. https://doi.org/10.1007/s004390050400

Stevens, G., van Beukering, J., Jenkins, T., Ramsay, M., 1995. An intragenic deletion of the P gene is the common mutation causing tyrosinase-positive oculocutaneous albinism in southern African Negroids. Am J Hum Genet 56, 586–591.

Thomas, M.G., Zippin, J., Brooks, B.P., 1993. Oculocutaneous Albinism and Ocular Albinism Overview, in: Adam, M.P., Feldman, J., Mirzaa, G.M., Pagon, R.A., Wallace, S.E., Amemiya, A. (Eds.), GeneReviews®. University of Washington, Seattle, Seattle (WA).

Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14, 178–192. https://doi.org/10.1093/bib/bbs017

Wei, A., Wang, Yu, Long, Y., Wang, Yi, Guo, X., Zhou, Z., et al., 2010. A Comprehensive Analysis Reveals Mutational Spectra and Common Alleles in Chinese Patients with Oculocutaneous Albinism. J Invest Dermatol 130, 716–724. https://doi.org/10.1038/jid.2009.339

Witkop, C.J., 1989. Albinism. Clinics in Dermatology 7, 80–91. https://doi.org/10.1016/0738-081X(89)90059-X

Yao, R., Zhang, C., Yu, T., Li, N., Hu, X., Wang, X., et al., 2017. Evaluation of three read-depth based CNV detection tools using whole-exome sequencing data. Molecular Cytogenetics 10, 30. https://doi.org/10.1186/s13039-017-0333-5

Yuan, N., Jia, P., 2024. Comprehensive assessment of long-read sequencing platforms and calling algorithms for detection of copy number variation. Briefings in Bioinformatics 25, bbae441. https://doi.org/10.1093/bib/bbae441

Online resources

GnomAD:

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

The Genome Aggregation Database (gnomAD). Available at: https://gnomad.broadinstitute.org/ (accessed 28 November 2024)

OMIM:

Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. MIM Number: 203200. 25 October 2023. World Wide Web URL: https://omim.org/entry/203200 (accessed 28 November 2024)

Varsome:

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C.E., Albarca Aguilera, M., Meyer, R., et al., 2019. VarSome: the human genomic variant search engine. Bioinformatics 35, 1978–1980. https://doi.org/10.1093/bioinformatics/bty897

VarSome: The human genomic variant search engine. Available at: https://varsome.com/ (accessed 28 November 2024)

Appendices

Appendix A: Supplementary material

- **Supplementary table 1**: Coverage of identified variants in each patient
- **Supplementary figure 1**: IGV visualisation of the OCA2: NM_000275.3: c.1239+2T>A variant in patient 11
- **Supplementary figure 2**: IGV visualisation of the TYRP1: NM_000550.3: c. 291G>C variant in patient 16
- **Supplementary figure 3**: IGV visualisation of the OCA2: NM_000275.3: c.1007C>T variant in patient 10 and 13
- **Supplementary figure 4**: IGV visualisation of the TYRP1: NM_000550.3: c.1557T>G variant in patient 13
- **Supplementary figure 5**: IGV visualisation of the MC1R: NM_002386.4: c.586T>C variant in patient 11
- **Supplementary table 2**: SpliceAI scores and their functional implications for the *OCA2:* c.1239+2T>A variant
- **Supplementary document 1**: ACMG summarised guidelines

Appendix B: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M240749

Appendix C: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M210203

Appendix D: Approved research protocol

Appendix E: Plagiarism Declaration

Appendix F: Turnitin Report

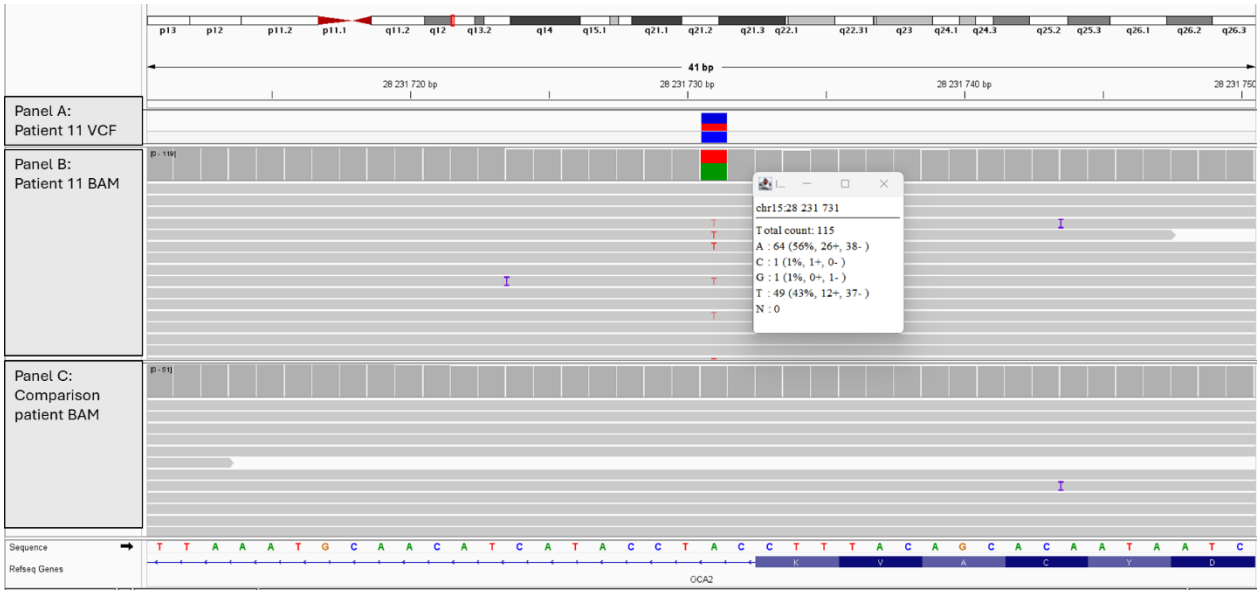
Appendix G: *Frontiers in Genetics* Author guidelines

Appendix A: Supplementary material

Supplementary table 1: Coverage of identified variants in each patient

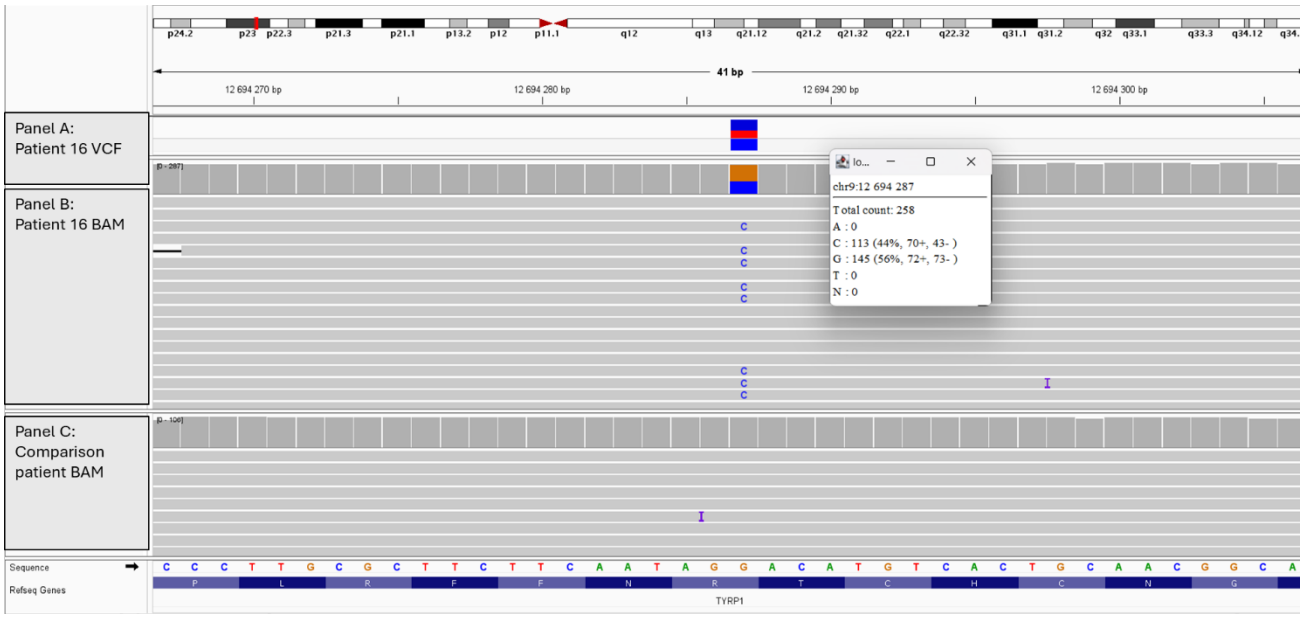
Patient ID	Gene	Variant	Coverage
10	*OCA2*	NM_000275.3: c.1007C>T	93x
11	*MC1R*	NM_002386.4: c.586T>C	44x
11	*OCA2*	NM_000275.3: c.1239+2T>A	115x
13	*OCA2*	NM_000275.3: c.1007C>T	88x
13	*TYRP1*	NM_000550.3: c.1557T>G	297x
16	*TYRP1*	NM_000550.3: c.291G>C	258x

IGV plots:



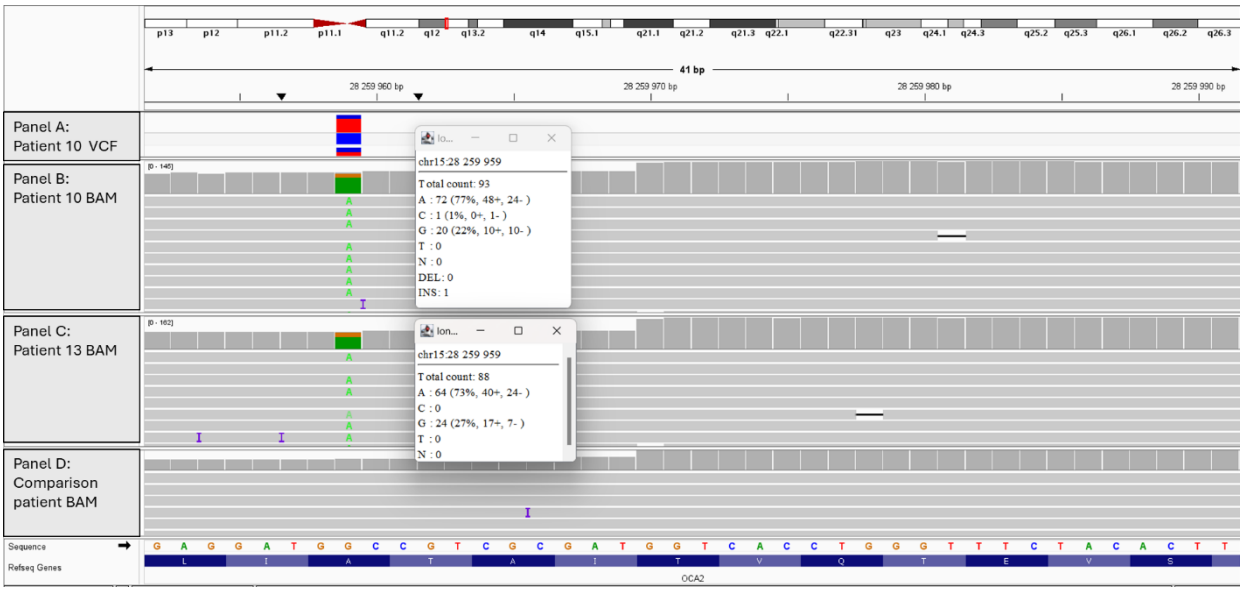
Supplementary figure 1*: IGV visualisation of the OCA2: NM_000275.3: c.1239+2T>A variant in patient 11.*

The variant is located in chromosome 15:28231731and appears to be heterozygous, with a total read depth of 115. Panel A and B displays the VCF and BAM files of patient 11, panel C displays the random comparison patient file (non-OCA2 patient).



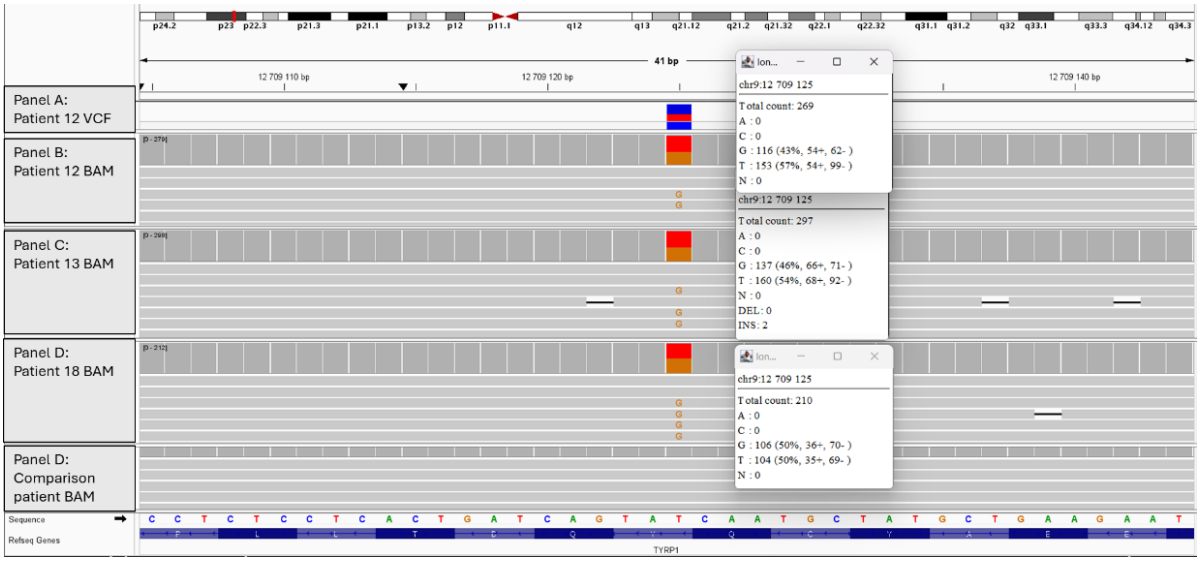
Supplementary figure 2: *IGV visualisation of the TYRP1: NM_000550.3: c. 291G>C variant in patient 16.*

The variant is located in chromosome 9:12694287 and appears to be heterozygous, with a total read depth of 258. Panel A and B displays the VCF and BAM files of the patient and panel C displays the random comparison patient file (non-OCA2 patient).



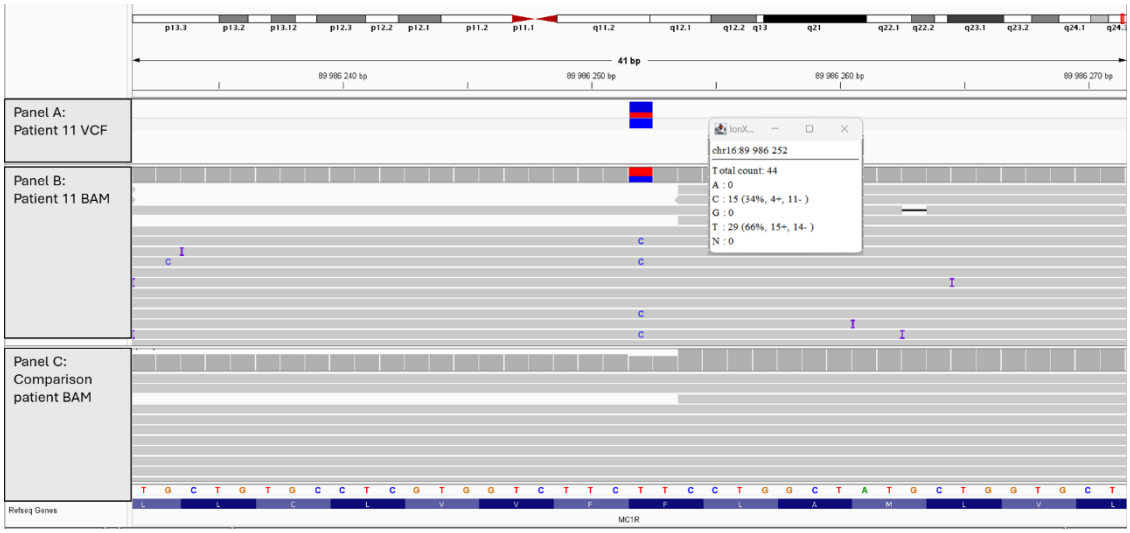
Supplementary figure 3: *IGV visualisation of the OCA2: NM_000275.3: c.1007C>T variant in patient 10 and 13.*

The variant is located in chromosome 15:28259959 and appears to be heterozygous, with a total read depth of 93 (P10) and 88 (P13). Panel A and B displays the VCF and BAM files of patient 10, panel C displays the BAM file of patient 13 and panel D displays the random comparison patient file (non-OCA2 patient).



Supplementary figure 2: *IGV visualisation of the TYRP1: NM_000550.3: c.1557T>G variant in patient 13.*

The variant is located in chromosome 9:12709125 and appears to be heterozygous, with a total read depth of 115. Panel A and B displays the VCF and BAM files of patient 11, panel C displays the random comparison patient file (non-OCA2 patient).



Supplementary figure 3*: IGV visualisation of the MC1R: NM_002386.4: c.586T>C variant in patient 11.*

The variant is located in chromosome 16:89986252 and appears to be heterozygous, with a total read depth of 44. Panel A and B displays the VCF and BAM files of the patient and panel C displays the random comparison patient file (non-OCA2 patient).

Supplementary table 2*: SpliceAI scores and their functional implications for the OCA2: c.1239+2T>A variant*

SpliceAI scores		Explanation
Position for Donor Gain	-2	Indicates a new donor splice site could be gained 2 bases upstream of the variant.
Position for Donor Loss	2	Indicates the original donor splice site might be lost 2 bases downstream of the variant.
Score for Donor Gain	0.91	Suggests a high likelihood (91%) of a new donor splice site forming.
Score for Donor Loss	0.99	Indicates an extremely high likelihood (99%) that the original donor splice site is disrupted.

Supplementary document 1: ACMG summarised guidelines

© American College of Medical Genetics and Genomics
ACMG STANDARDS AND GUIDELINES | Genetics inMedicine

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

Sue Richards, PhD[1], Nazneen Aziz, PhD[2,16], Sherri Bale, PhD[3], David Bick, MD[4], Soma Das, PhD[5], Julie Gastier-Foster, PhD[6,7,8], Wayne W. Grody, MD, PhD[9,10,11], Madhuri Hegde, PhD[12], Elaine Lyon, PhD[13], Elaine Spector, PhD[14], Karl Voelkerding, MD[13] and Heidi L. Rehm, PhD[15]; on behalf of the ACMG Laboratory Quality Assurance Committee

The ACMG classification system differentiates sequence variants in Mendelian diseases regardless of the type of analysis used to identify them (Sanger or next-generation sequencing).

Variants are divided into 5 classes in accordance with IARC recommendations (Plon et al.; *Hum Mutat* 2008):

Class 1	benign
Class 2	likely benign
Class 3	variant of uncertain significance (VUS)
Class 4	likely pathogenic
Class 5	pathogenic

The following 3 tables are based on the primary literature and provide a schematic overview of the ACMG classification system. For detailed information we recommend reading the publication in its entirety.

Table 1lists the criteria for classifying a variant as pathogenic. Table 2 shows the criteria for classifying a variant as benign. In order to make an unambiguous classification of detected variants, these criteria are combined according to the rules laid out in Table 3.

Criteria for classifying pathogenic variants (Table 1)

Evidence of pathogenicity		Category
Very strong	PVS1	**Null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease.** *Caveats:* • *Beware of genes where LOF is not a known disease mechanism (e.g., GFAP, MYH7)* • *Use caution interpreting LOF variants at the extreme 3' end of a gene* • *Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact* • *Use caution in the presence of multiple transcripts*
Strong	PS1	**Same amino acid change as a previously established pathogenic variant regardless of nucleotide change** • *Example: Val→Leu caused by either G>C or G>T in the same codon* • *Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level*
	PS2	**De novo (both maternity and paternity confirmed) in a patient with the disease and no family history** *Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, and so on, can contribute to non maternity.*
	PS3	**Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product** *Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well established.*
	PS4	**The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls** *Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0.* *See the article for detailed guidance.* *Note 2: In instances of very rare variants where case–control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.*
Moderate	PM1	**Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation.**
	PM2	**Absent from controls (or at extremely low frequency if recessive) (Table 6) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium** • *Caveat: Population data for insertions/deletions may be poorly called by next-generation sequencing.*
	PM3	**For recessive disorders, detected in trans with a pathogenic variant** *Note: This requires testing of parents (or offspring) to determine phase.*
	PM4	**Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants**
	PM5	**Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before** • *Example: Arg156His is pathogenic; now you observe Arg156Cys* • *Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level.*
	PM6	**Assumed de novo, but without confirmation of paternity and maternity**

Supporting	PP1	**Co segregation with disease in multiple affected family members in a gene definitively known to cause the disease** *Note: May be used as stronger evidence with increasing segregation data*
	PP2	**Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease**
	PP3	**Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.)** • *Caveat: Because many in-silico algorithms use the same or very similar input for their predictions, each algorithm should not be counted as an independent criterion. PP3 can be used only once in any evaluation of a variant.*
	PP4	**Patient's phenotype or family history is highly specific for a disease with a single genetic etiology**
	PP5	**Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation**

Criteria for classifying benign variants (Table 2)

Evidence of benign impact		Category
Stand Alone	BA1	**Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium**

Strong	BS1	**Allele frequency is greater than expected for disorder (see Table 6)**
	BS2	**Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age**
	BS3	**Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing**
	BS4	**Lack of segregation in affected members of a family** • Caveat: The presence of phenocopies for common phenotypes (i.e., cancer, epilepsy) can mimic lack of segregation among affected individuals. Also, families may have more than one pathogenic variant contributing to an autosomal dominant disorder, further confounding an apparent lack of segregation.

Supporting	BP1	**Missense variant in a gene for which primarily truncating variants are known to cause disease**
	BP2	**Observed in trans with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern**
	BP3	**In-frame deletions/insertions in a repetitive region without a known function**
	BP4	**Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)** • Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm cannot be counted as an independent criterion. BP4 can be used only once in any evaluation of a variant.
	BP5	**Variant found in a case with an alternate molecular basis for disease**
	BP6	**Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation**
	BP7	**A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved**

Rules for combining criteria to classify sequence variants
(Table 3)

Pathogenic	**(i) 1 Very strong (PVS1) AND** *a) ≥1 Strong (PS1–PS4) OR* *b) ≥2 Moderate (PM1–PM6) OR* *c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) OR* *d) ≥2 Supporting (PP1–PP5)*
	(ii) ≥2 Strong (PS1–PS4) OR
	(iii) 1 Strong (PS1–PS4) AND *a) ≥3 Moderate (PM1–PM6) OR* *b) 2 Moderate (PM1–PM6) AND ≥2 Supporting (PP1–PP5) OR* *c) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)*
Likely pathogenic	**(i) 1 Very strong (PVS1) AND 1 moderate (PM1–PM6) OR**
	(ii) 1 Strong (PS1–PS4) AND 1–2 moderate (PM1–PM6) OR
	(iii) 1 Strong (PS1–PS4) AND ≥2 supporting (PP1–PP5) OR
	(iv) ≥3 Moderate (PM1–PM6) OR
	(v) 2 Moderate (PM1–PM6) AND ≥2 supporting (PP1–PP5) OR
	(vi) 1 Moderate (PM1–PM6) AND ≥4 supporting (PP1–PP5)
Benign	**(i) 1 Stand-alone (BA1) OR**
	(ii) ≥2 Strong (BS1–BS4)
Likely benign	**(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) OR**
	(ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	**(i) Other criteria shown above are not met OR**
	(ii) the criteria for benign and pathogenic are contradictory

Appendix B: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M240749

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

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

TO: Ms P Moodley
School of Pathology
Division of Human Genetics
Faculty of Health Sciences

E-mail: 2111913@students.wits.ac.za

CC: Supervisor: Drs T Ngcungcu, P Mpangase and R Kerr
Thandiswa.Ngcungcu@wits.ac.za
and <HREC-Medical Research Office@wits.ac.za>

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

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

DATE: 7 August 2024

REF: R14/49

PROTOCOL NO: **M240749** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)

PROJECT TITLE: *Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant*

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

MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

R49 Ms P Moodley

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M240749

NAME: (Principal Investigator)	Ms P Moodley **DEGREE:** MSc (Med)
DEPARTMENT:	School of Pathology Division of Human Genetics Faculty of Health Sciences
TITLE:	*Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant*
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M21/02/03
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office <Nicoleen.Potgieter@wits.ac.za>
SUPERVISORS:	Drs T Ngcungcu, P Mpangase and R Kerr
APPROVED BY:	Professor P Ruff, Chairperson, HREC (Medical)

DATE OF APPROVAL: 7 August 2024 **EXPIRY DATE:** 6 August 2029

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

DECLARATION OF INVESTIGATORS

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

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

Signature of Principal Investigator Date

Appendix C: Human Research Ethics Committee (Medical), University of the Witwatersrand, Clearance Certificate No. M210203

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

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

TO:	Dr T Ngcungcu School of Pathoogy Department of Human Genetics National Health Laboratory Service E-mail: Thandiswa.Ngcungcu@wits.ac.za
CC:	Supervisor: Not applicable <> and <HREC-Medical Research Office@wits.ac.za>
FROM:	Mr Iain Burns Human Research Ethics Committee (Medical) Tel: 011 717 1252 E-mail: Iain.Burns@wits.ac.za
DATE:	2021/08/10
REF:	R14/49
PROTOCOL NO:	**M210203** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)
PROJECT TITLE:	*The identification of causal mutations of classic oculocutaneous albinism type 2 and brown albinism in individuals of African descent*

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



MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R49 Dr T Ngcungcu

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M210203

NAME: **(Principal Investigator)**	Dr T Ngcungcu
DEPARTMENT:	School of Pathoogy Department of Human Genetics National Health Laboratory Service
PROJECT TITLE:	*The identification of causal mutations of classic oculocutaneous albinism type 2 and brown albinism in individuals of African descent*
DATE CONSIDERED:	2021/02/26
DECISION:	Approved unconditionally
CONDITIONS:	
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Not applicable
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2021/08/10

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.

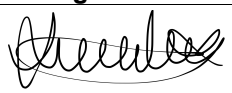
I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. **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 **February** and therefore reports and re-certification will be due in the month of **February** each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator _____ Date _____

Appendix D: Approved research protocol

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

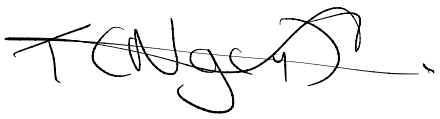
CANDIDATE'S SURNAME: [Please print] **Moodley**	FIRST NAME/S: **Priya**	STUDENT NUMBER: **2111913**
CURRENT QUALIFICATIONS: **BHSc (Hons) Human Genetics**		
TEL: **N/A**	CELL: **072 999 1074**	E-MAIL: **2111913@students.wits.ac.za** FAX: **N/A**
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: **MSc (Med) Genomic Medicine**		
PART-TIME OR FULL-TIME: **Full-time**		
FIRST REGISTERED FOR THIS DEGREE:	TERM : **1**	YEAR: **2024**
DEPARTMENT: **Division of Human Genetics, School of Pathology, University of the Witwatersrand**		
TITLE OF PROPOSED RESEARCH: **Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant**		
CANDIDATE'S SIGNATURE: 		DATE: **14/06/2024**
SUPERVISOR 1 (NAME & SURNAME): **Dr Thandiswa Ngcungcu**		% Supervision = **50**
SUPERVISOR'S QUALIFICATIONS: **PhD**		
SUPERVISOR'S DEPARTMENT: **Division of Human Genetics, University of the Witwatersrand**		
SUPERVISOR'S ADDRESS / TEL / E-MAIL: **Thandiswa.Ngcungcu@wits.ac.za**		
SUPERVISOR 2 (NAME & SURNAME): **Dr Robyn Kerr**		% Supervision = **25**
SUPERVISOR'S QUALIFICATIONS: **PhD**		
SUPERVISOR'S ADDRESS / TEL / E-MAIL: **Robyn.Kerr@wits.ac.za**		
SUPERVISOR 3 (NAME & SURNAME): **Dr Phelelani T. Mpangase**		% Supervision = **25**
SUPERVISOR'S QUALIFICATIONS: **PhD**		
SUPERVISOR'S ADDRESS / TEL / E-MAIL: **Phelelani.Mpangase@wits.ac.za**		

SYNOPSIS OF RESEARCH:*(Brief summary of proposed research project; between 200-300 words only; with sub-headings: an introduction and justification for study, aim/s, proposed methodology and expected outcome/s)*
[Use reverse side of this page if more space is required]

See next page.

WITS ETHICS NOT REQUIRED: WITS ETHICS PENDING: WITS ETHICS APPROVED: (circle appropriate symbol)*	Yes (No) (Yes) No Yes (No)	IF Y SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE:

***Please note the final human ethics clearance certificate or animal ethics certificate must be available prior to starting research**

As supervisor/s, I/we confirm that I have read the protocol which has been submitted for assessment.		
SIGNATURE OF SUPERVISOR/S:		**Dr Thandiswa Ngcungcu**
		**Dr Robyn Kerr**
		**Dr Phelelani T. Mpangase**
SIGNATURE PG OFFICE STAFF ……..	REGISTERED YES….. NO…..	STAMP

SYNOPSIS OF RESEARCH CONTINUED

Introduction:

Albinism is a group of rare, inherited pigmentation disorders characterized by hypopigmentation of the skin and impaired vision. The most common form of albinism is oculocutaneous albinism type 2 (OCA2) which has a high prevalence in African populations. The most common cause of OCA2 in African individuals is a 2.7kb intragenic deletion variant in the *OCA2* gene, and it accounts for 78% of all pathogenic variants in this population, leaving the variant unknown for 22% of alleles. Identifying the remaining pathogenic variants will be the focus of this study.

Aim:

The aim of this study is to identify pathogenic variants in African individuals with OCA2 that are negative for the common African *OCA2* 2.7 kb deletion variant.

Methods:

Whole exome sequencing (WES) has been performed on samples as part of the parent study (M210203). Whole exome sequencing data of ten individuals who were clinically diagnosed with OCA2 and sent to the Diagnostic Laboratory, Division of Human Genetics, NHLS for genetic testing, but who tested homozygous negative for the 2.7kb deletion, will be analysed to identify *OCA2* pathogenic variants. This will be done through bioinformatics analysis of WES data and will include variant annotation, prioritization, visualization and interpretation. Copy number variant analysis will also be performed using ThermoFisher's Ion Reporter software and other open source tools. Prioritised variants will be classified according to the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP) guidelines.

Expected outcomes:

Possible identification of any other clinically significant variants, in *OCA2* as well as other skin pigmentation related genes, which contribute to an OCA2 phenotype in individuals that have tested negative for the common *OCA2* 2.7kb deletion variant.



Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant

Research Proposal

PRIYA MOODLEY (2111913)

Master of Science in Medicine in the field of Genomic Medicine (MFAGENM60)

Division of Human Genetics

12th April 2024

Supervisors:

1. **Dr Thandiswa G. Ngcungcu**

   PhD (Wits), Lecturer, Division of Human Genetics, University of the Witwatersrand
2. **Dr Robyn A. Kerr**

   PhD (Wits), Lecturer, Division of Human Genetics, University of the Witwatersrand
3. **Dr Phelelani T. Mpangase**

   PhD (Bioinformatics) Wits, Bioinformaticist, Sydney Brenner Institute of Molecular Bioscience, University of the Witwatersrand

Table of Contents

1. Background literature analysis and critique 1

1.1. Introduction 1

1.2. Oculocutaneous albinism type 2: clinical and genetic aspects 1

1.2.1. Clinical presentation of oculocutaneous albinism type 2 2

1.2.2. Genetic aetiology of OCA2 2

1.3. Genetic testing for albinism 3

2. Rationale 4

3. Aim of the study 5

4. Study objectives 5

5. Methods 5

5.1. Study Participants 5

5.2. Data analysis 6

6. Ethics 8

7. Timing 9

8. Funding 9

9. Limitations 9

10. References 10

Appendix A 13

List of abbreviations

Abbreviations	Definitions
ACMG-AMP	American College of Medical Genetics and Genomics and the Association for Molecular Pathology
BAM	Binary alignment map
CNV	Copy number variant
IGV	Integrative genomics viewer
MAF	Minor allele frequency
MRC-SIR	Medical Research Council - Self-Initiated Research award
NHLS	National health laboratory service
NGS	Next generation sequencing
OA	Ocular albinism
OCA	Oculocutaneous albinism
OCA2	Oculocutaneous albinism type 2
OMIM	Online Mendelian Inheritance in Man
PCR	Polymerase chain reaction
PI	Principal investigator
SNV	Single nucleotide variant
VCF	Variant call format
VEP	Variant effect predictor
VUS	Variant of uncertain significance
WES	Whole exome sequencing

1. Background literature analysis and critique

1.1. Introduction

Albinism is a group of rare, inherited pigmentation disorders characterized by hypopigmentation and impaired vision (Kromberg and Manga, 2018). Albinism exhibits an autosomal recessive pattern of inheritance. There are two main types of albinism, ocular albinism (OA) and oculocutaneous albinism (OCA). OCA is characterized by a lack of melanin pigment (hypopigmentation) of the skin, hair, and eyes from birth (King et al., 1995).

The worldwide prevalence for all forms of albinism varies greatly, but is estimated at approximately 1/17000 (Witkop et al., 1983). There are nine types of OCA, and the other types are caused by variants at different loci (Kromberg et al., 2012). The most prevalent form of OCA worldwide is oculocutaneous albinism type 2 (OCA2), as it is the type of OCA affecting individuals of African descent. The prevalence for OCA2 is 1/36000 for individuals of European descent and 1/3900 for individuals of African descent (Kromberg et al., 2020).

OCA is caused by defects in the melanin biosynthesis pathway. This pathway comprises of a series of biochemical reactions that lead to melanin production, which is the pigment responsible for skin, hair, and eye pigmentation (Riley, 1997). There are two melanin precursors: eumelanin, which creates a brown or black coloured pigment, and pheomelanin, which creates a red or yellow coloured pigment (Riley, 1997). Eumelanin is also ultraviolet absorbent, thus protects the skin from the sun's radiation (Thody et al., 1991). Pheomelanin is photo-chemically unstable and after sun exposure can promote carcinogenesis (Thody et al., 1991). An imbalance of these pigments in the body will result in disorders like OCA2. OCA2 is differentiated from the other types of OCA because pheomelanin is present and accumulates over time (Thody et al., 1991).

1.2.Oculocutaneous albinism type 2: clinical and genetic aspects

Oculocutaneous albinism type 2, is the most common form of albinism. It is especially prevalent amongst individuals of African descent. The condition most commonly results from variants in

the *OCA2* gene, affecting melanin production. This lack of melanin leads to very light skin and hair and can cause various vision problems (Kromberg and Manga, 2018). Exploring the underlying variants that cause OCA2 will enable researchers and healthcare providers to provide better diagnosis, management, and support to individuals affected by this condition, ultimately improving their quality of life.

1.2.1. Clinical presentation of oculocutaneous albinism type 2

Individuals with OCA2 present with creamy white skin, impaired vision, nystagmus , strabismus and photophobia (The National Institutes of Health, 2024; Witkop, 1989). Some affected individuals present with skin patches called ephelides (appear like large freckles) and hair that appears straw-coloured (yellow – blonde) or even light brown as some pigment accumulates over time (Kromberg et al., 1989). In individuals with OCA2, eumelanin is absent from skin, hair, and eyes however pheomelanin is present in smaller quantities and accumulates with age. This can lead to an increased risk of skin damage and skin cancers such as melanoma for OCA2 individuals (Kromberg et al., 1989).

1.2.2. Genetic aetiology of OCA2

Initially, it was thought albinism resulted from various pathogenic variants at a single genetic locus, but investigations disproved this hypothesis, revealing locus heterogeneity in OCA. OCA2 is commonly associated with variants in the *OCA2* gene on chromosome 15q (OMIM 203200). OCA2 exhibits allelic heterogeneity in black individuals from southern Africa (Kedda et al., 1994).

In 1994, Durham-Pierre et al discovered the 2.7kb deletion in the *OCA2* gene, which is the most common pathogenic variant in populations of African descent. The study found that the deletion was an insertional frameshift deletion, which removes a single exon of the *OCA2* gene, resulting in a truncated and non-functional gene (Durham-Pierre et al., 1994).

There are various genes involved in skin pigmentation. D'Arcy et al have compiled a comprehensive list of over 200 genes that are associated with pigmentation and pigmentation

disorders (D'Arcy et al., 2022). A shortlist of these genes associated with albinism has been included in appendix A (pg. 13).

An important concept to keep in mind is epistasis (gene-gene interactions). Epistasis within the melanin biosynthesis pathway involves multiple genes such as *TYR, TYRP1, TYRP2, SLC24A5, MC1R, OCA2* and *PMEL* (Chang et al., 2024). Changes to any of the genes in this pathway could affect melanin production and result in disease. A more specific example is the interaction between the *HERC2* and *OCA2* genes. *HERC2* is said to have an influence of *OCA2* because it has an effect on the transcriptional activity of *OCA2*, being able to regulate the gene's expression. Single nucleotide variants (SNVs) in this region of *HERC2* which controls *OCA2* have also shown to decrease melanin production (as reviewed by Brancato et al., 2023).

With regards to albinism specifically, many studies have been carried out in other populations such as the Japanese and Polynesian population groups, however studying albinism in African populations will provide more insight into the disease for the population it affects most (Kromberg and Manga, 2018).

The most common variant that causes OCA2 in individuals of African descent is the 2.7kb deletion variant that includes exon 7 and it accounts for 78% of all pathogenic variants in this population, leaving the variant unknown for 22% of alleles (Stevens et al., 1995). This means that several individuals with OCA2 are either heterozygous for the deletion, or do not have it at all. Investigation into individuals who are negative 2.7kb deletion individuals can provide great insight into the disease for these individuals as they may have variants in other genes or may even have a different condition entirely, which manifests with a similar phenotypic presentation as OCA2.

1.3. Genetic testing for albinism

There are various molecular genetic tests available for OCA2 such as targeted variant analysis, deletion and duplication analysis, exon mutation scanning, exon sequence analysis and sequence analysis of the entire coding region through techniques like next-generation sequencing (NGS).

Generally, a molecular test is not needed to clinically diagnose OCA2, however molecular tests can be done for research purposes, prenatal testing or to identify a specific mutation (especially within families that have the same phenotype) (The National Institutes of Health, 2022). Currently, the only testing available through the National Health Laboratory Service (NHLS), Johannesburg in South Africa is the gap-polymerase chain reaction (PCR) test developed by Durham-Pierre et al for the common 2.7kb deletion, which has been used since 1996 as it is efficient and affordable (Kerr et al., 2000). Follow-up testing is not available due to the resource-constrained setting in South Africa.

Whole exome sequencing (WES) is a technique used to sequence the protein coding regions of the genome. This technique is useful because most disease-causing known mutations are found in the coding regions and it allows for detection of variants in any gene rather than a select few (Meienberg et al., 2016). WES is an appropriate technique to use for investigating variants in *OCA2* and other skin pigmentation related genes. The *OCA2* gene is quite large, having 24 exons and sequencing all the exons of the *OCA2* gene as well as the other related genes will require an efficient and high throughput technique such as WES (Hawkes et al., 2013). This also makes WES a cost-effective choice. WES can detect single nucleotide variants (SNVs) and small insertions or deletions with high accuracy. Moreover, it is also useful for patients whose initial testing did not yield a diagnosis (Meienberg et al., 2016). This is useful in identifying novel genetic variants, such as identifying any variants in individuals of African descent with OCA2 who were negative for the common African 2.7kb deletion variant (Meienberg et al., 2016).

2. Rationale

This sub-study is part of a larger OCA2 study that is investigating individuals who are either heterozygous for the 2.7kb deletion or do not have the deletion at all. Previously a project was done looking at the heterozygous individuals and now, for this current project, the individuals who

did not have the 2.7kb deletion will be investigated. The findings will inform future OCA testing and contribute to the understanding of the *OCA2* variant profile.

3. Aim of the study

The aim of this study is to identify pathogenic variants in African individuals with OCA2 that are negative for the common African *OCA2* 2.7 kb deletion variant.

4. Study objectives

1. Identify and analyse variants in the *OCA2* gene using previously generated WES data.

    1.1. **Single nucleotide variant (SNV) analysis:** annotation of variants using Ensembl's Variant Effect Predictor (VEP) tool, then filtering and prioritisation and lastly classification according to the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) guidelines for SNVs.

    1.2. **Copy number variant (CNV) analysis:** annotation, filtering and prioritisation of variants using Thermofisher's Ion Reporter software, then classification according to the ACMG-AMP guidelines for CNVs.

2. Identify and analyse other genes associated with OCA in patients with no clinically significant *OCA2* variants using previously generated WES data as in objective 1.

5. Methods

5.1. Study Participants

The study participants consist of ten individuals of African descent with OCA2 who were found to be negative for the 2.7kb deletion in the *OCA2* gene. The WES data has already been generated under the parent study (larger OCA2 study) and is available for analysis.

5.2. Data analysis

The ten samples have been genotyped by gap PCR to confirm the presence of the 2.7kb deletion (as part of the parent study). Gap PCR to genotype the deletion will be completed for each sample in the parent study. PCR products will be visualised on a 3% agarose gel, and these will be used to determine the 2.7kb deletion genotypes. This is an example of the gel image that will be analysed:

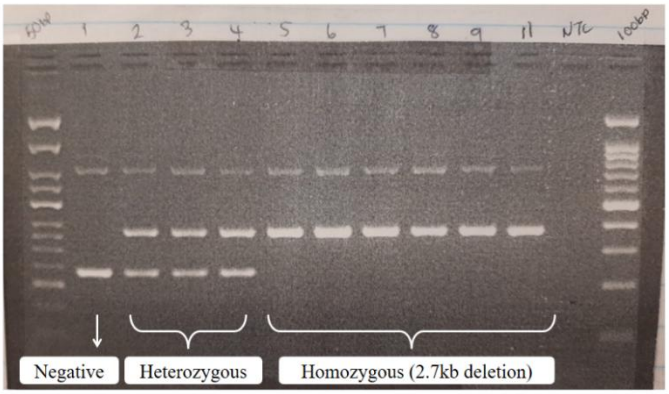


Figure 1: Gel electrophoresis image of *OCA2* 2.7kb deletion genotypes generated by a gap-PCR (Ngcungcu, 2023)

Whole exome sequencing data has already been generated using the Ion Torrent™ Ion S5 and the data is available in the form of a variant call format (VCF) file, binary map alignment (BAM) file and index file. The raw sequencing data generated from WES will be analysed as indicated in figure 2:

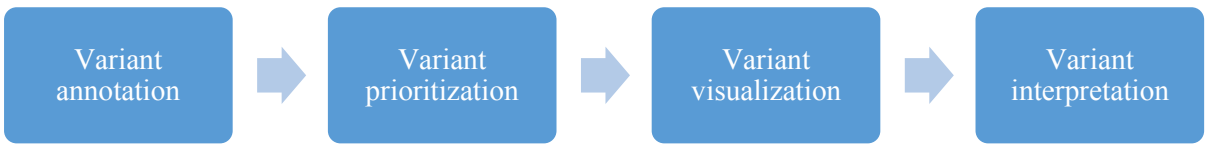


Figure 2: Flow diagram illustrating variant analysis

Variants will be annotated using Ensembl's VEP tool. The VEP tool is a bioinformatics tool that is used to analyse and annotate sequencing data. VEP determines the effect of variants on genes, transcripts, protein sequences and regulatory regions (McLaren et al., 2016). Variant annotation involves assigning functional information and significance to variants in DNA, such as the type of

variant and the level of pathogenicity of the variant (McCarthy et al., 2014). Information such as gene symbols, accession numbers, population frequency data and in-silico prediction scores are also provided. Then, the variants will be filtered and prioritised based on functionality. Frameshift, missense, non-synonymous and splice site variants with a low minor allele frequency (MAF<5%) will be prioritised.

Next, a quality assessment will be done, using the Integrative Genomics Viewer (IGV) tool, which is used to visualize large genomic datasets. Variants in the human genome can be visualized using IGV since it can manage larger sets of data (Thorvaldsdóttir et al., 2013). This tool was used to assess the depth of coverage, confirm zygosity, and exclude artefacts for each prioritised variant. If the variant cannot be located it is most likely unreliable, a false call or there may be low coverage (reads < 100) of that region on IGV.

Prioritised variants will be classified based on the ACMG-AMP guidelines for variant interrogation. The ACMG-AMP guidelines provide a framework for classifying genetic variants into different categories based on evidence such as population data, computational predictions, functional studies, segregation with disease, de novo occurrence, and allelic data. After combining the evidence and assessing their strength, variants are classified into five categories: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign. This helps geneticists, clinicians, and researchers determine whether a genetic variant might be responsible for a disease or condition (Richards et al., 2015). For the purpose of this project, the following clinically significant classifications will be assigned: pathogenic, likely pathogenic and variant of uncertain significance (VUS).

Copy number variant analysis will also be performed using ThermoFisher's Ion Reporter software. Ion Reporter Software detects copy number variants (CNVs) by first aligning sequencing reads to a reference genome and calculating normalized read depth. It then uses segmentation algorithms and statistical analysis to identify regions with significant changes in read depth, indicating

potential CNVs. Finally, the software annotates and visualizes these CNVs, providing detailed reports on their genomic locations and potential clinical implications (Thermo Fisher Scientific, 2020). Copy number variants with structural consequences (predicted loss-of-function (LoF) and copy gain variants), rare or novel CNVs (based on frequency on DGV) will be prioritised and classified according to ACMG-AMP guidelines.

If no variants meet the filtering criteria within the *OCA2* gene, the same methods will be performed for other skin pigmentation related genes, that cause other types of OCA. There are eight other genes, associated with OCA2, that will be investigated (*TYR, MC1R, TYRP1, SLC45A2, SLC24A5, C10ORF11, LRMDA, DCT*). Variants will be identified, analysed and assessed according to the above criteria.

6. Ethics

This research project is a sub-study that falls under the "The identification of causal mutations of classic oculocutaneous albinism type 2 and brown albinism in individuals of African descent" parent study (PI - Dr Thandiswa Ngcungcu, ethics number M210203). This MSc research project is currently under consideration by the Wits Human Research Ethics Committee. Ethics application was submitted on 15th March 2024. This project will begin once ethics has been approved, possibly beginning of June 2024.

7. Timing

	2024									
	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
Literature Review										
Ethics Application										
Protocol Preparation										
Protocol Assessment										
Data Analysis										
Write Up										

8. Funding

This research project is a sub-study that falls under a parent study that received funding for the resources and services required from the Medical Research Council Self-Initiated Research (MRC-SIR) award. The WES data required for the current MSc research project was generated from this parent study, therefore, no funding is required for the analysis of the sub-study.

9. Limitations

This project has some limitations regarding the sample size, WES and variants of uncertain significance. The small sample size (10 individuals) may hinder the ability to detect common variants in negative deletion individuals and hinder the generalizability of the findings to broader populations. Whole exome sequencing only targets the protein-coding regions of the genome, leaving out non-coding regions, which overlaps with many regulatory elements. Structural variations that could be relevant for OCA2 and skin pigmentation may be missed. Certain covered by WES regions may also have poor coverage and read depth. Data from WES may identify many VUS, which can complicate clinical interpretation. Historically, African populations have been underrepresented in genomic studies and databases. African populations exhibit high levels of genetic diversity, which can lead to a greater number of rare variants and potentially unique genetic variants not being well characterised in existing databases. Variants identified in African populations may be classified as VUS due to insufficient data on their frequency or functional

consequences. Without robust population-specific data, it can be challenging to determine the clinical significance of these variants.

10. References

- Chang, Y., Wu, S., Li, J., Bao, H., Wu, C. (2024). Identification of Candidate Genes for Red-Eyed (Albinism) Domestic Guppies Using Genomic and Transcriptomic Analyses. *International Journal of Molecular Sciences,* 25, 2175. https://doi.org/10.3390/ijms25042175
- Durham-Pierre, D., Gardner, J.M., Nakatsu, Y., King, R.A., Francke, U., Ching, A., et al. (1994). African origin of an intragenic deletion of the human P gene in tyrosinase positive oculocutaneous albinism. *Nature Genetics,* 7, 176–179. https://doi.org/10.1038/ng0694-176
- Hawkes, J.E., Cassidy, P.B., Manga, P., Boissy, R.E., Goldgar, D., Cannon-Albright, L., et al. (2013). Report of a Novel OCA2 Gene Mutation and an Investigation of Two OCA2 Variants on Melanoma Predisposition in a Familial Melanoma Pedigree. *Journal of Dermatological Science,* 69, 30–37. https://doi.org/10.1016/j.jdermsci.2012.09.016
- Kedda, M.A., Stevens, G., Manga, P., Viljoen, C., Jenkins, T., Ramsay, M. (1994). The tyrosinase-positive oculocutaneous albinism gene shows locus homogeneity on chromosome 15q11-q13 and evidence of multiple mutations in southern African negroids. T*he American Journal of Human Genetics*, 54, 1078–1084.
- King, R.A., Hearing, V.J., Creel, D.J., Oetting, W.S. (1995). Albinism, in: Valle, D.L., Antonarakis, S., Ballabio, A., Beaudet, A.L., Mitchell, G.A. (Eds.), The Metabolic and Molecular Bases of Inherited Disease. New York, NY: McGraw-Hill Education.

- Kromberg, J., Manga, P. (2018). Albinism in Africa. Academic Press, London, United Kingdom.
- Kromberg, J.G.R., Bothwell, J., Kidson, S.H., Manga, P., Kerr, R., Jenkins, T. (2012). Types of albinism in the black southern africa population. *East African Medical Journal*, 89, 20–27.
- Kromberg, J.G.R., Castle, D., Zwane, E.M., Jenkins, T. (1989). Albinism and skin cancer in Southern Africa. *Clinical Genetics*, 36, 43–52. https://doi.org/10.1111/j.1399-0004.1989.tb03365.x
- Kromberg, J.G.R., Kerr, R., Manga, P. (2020). Children with oculocutaneous albinism in Africa: Characteristics, challenges and medical care. *South African Journal of Child Health,* 14, 50–54. https://doi.org/10.7196/SAJCH.2020. v14.i1.1608
- McCarthy, D.J., Humburg, P., Kanapin, A., Rivas, M.A., Gaulton, K., Cazier, J.-B., et al. (2014). Choice of transcripts and software has a large effect on variant annotation. *Genome Medicine*. 6, 26. https://doi.org/10.1186/gm543
- McLaren, W., Gil, L., Hunt, S.E., Riat, H.S., Ritchie, G.R.S., Thormann, A., et al. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*. 17, 122. https://doi.org/10.1186/s13059-016-0974-4
- Meienberg, J., Bruggmann, R., Oexle, K., Matyas, G. (2016). Clinical sequencing: is WGS the better WES? *Human Genetics*, 135, 359–362. https://doi.org/10.1007/s00439-015-1631-9
- Riley, P.A. (1997). Melanin. *The International Journal of Biochemistry & Cell Biology*, 29, 1235–1239. https://doi.org/10.1016/S1357-2725(97)00013-7
- Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., et al. (2015). Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetic Medicine*. 17, 405–424. https://doi.org/10.1038/gim.2015.30

- Thody, A.J., Higgins, E.M., Wakamatsu, K., Ito, S., Burchill, S.A., Marks, J.M. (1991). Pheomelanin as well as Eumelanin Is Present in Human Epidermis. *Journal of Investigative Dermatology,* 97, 340–344. https://doi.org/10.1111/1523-1747.ep12480680
- Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. 14, 178–192. https://doi.org/10.1093/bib/bbs017
- Witkop, C.J. (1989). Albinism. *Clinics in Dermatology*, 7, 80–91. https://doi.org/10.1016/0738-081X(89)90059-X
- Witkop, C.J., Quevedo, W.C., Fitzpatrick, T.P. (1983). Albinism and disorders of pigment metabolism, in: The Metabolic Basis of Inherited Disease. (pp 301-346). New York, NY: McGraw-Hill.

Online resources

- Ion Reporter™ Software 5.14 User Guide, Thermo Fisher Scientific, 2020.
- The National Institutes of Health. (2022). *Genetic Testing Registry - Oculocutaneous albinism*. https://www.ncbi.nlm.nih.gov/gtr/conditions/C0078918/ (accessed 3 March 2024).
- The National Institutes of Health. (2024). *Oculocutaneous albinism type 2*. URL https://rarediseases.info.nih.gov/diseases/4038/oculocutaneous-albinism-type-2 (accessed 29 February 2024).
- Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD. MIM Number: 203200. 25 October 2023. World Wide Web URL: https://omim.org/entry/203200 (accessed 3 March 2024)

Appendix A

Table of genes shown to be associated with albinism

OCA type	Gene symbol	Gene name
OCA1	*TYR*	Tyrosinase
OCA2	*OCA2*	Oculocutaneous albinism type II
	MC1R	Melanocortin 1 receptor
OCA3	*TYRP1*	Tyrosinase related protein 1
OCA4	*SLC45A2*	Solute carrier family 45, member 2
OCA5	*Region on chromosome 4q24 (14 genes)*	
OCA6	*SLC24A5*	Solute carrier family 24, member 5
OCA7	*C10ORF11/ LRMDA*	Leucine rich melanocyte differentiation associated
OCA8	*LRMDA*	Leucine rich melanocyte differentiation associated

Appendix E: Plagiarism Declaration

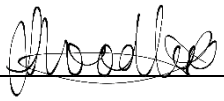
PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I **Priya Moodley** (Student number: **2111913**) am a student registered for the degree of **Master of Science in Medicine in the field of Genomic Medicine** in the academic year **2024**.

I hereby declare the following:

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

Signature:  Date: **29/11/2024**

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

2nd December 2024

FoHS Examination board

RE: High Turnitin Similarity Score in Ms. Priya Moodley's Research Report Submission

We are writing to provide a formal explanation for the Turnitin similarity score of 21% recorded for the thesis titled "Identification of pathogenic variants in African individuals with oculocutaneous albinism type 2 negative for the common 2.7 kb deletion variant" submitted by Ms. Priya Moodley as part of the requirements for her MSc in Genomic Medicine degree. While the similarity score above 15%, the higher score in this instance, can be fully justified and does not indicate plagiarism or academic misconduct.

The elevated score is primarily attributed to two factors:

1. Overlap with previously submitted protocols

Large portions of the Introduction and Methods sections of this report were drawn from a research protocol previously submitted to the university via Turnitin. Examples include unavoidable factual statements such as:

- "Albinism exhibits an autosomal recessive pattern of inheritance."
- "oculocutaneous albinism (OCA) is a group of rare."
- "oculocutaneous albinism type 2 (OCA2)"

These are standard scientific facts that are fundamental to the research topic and cannot be altered without compromising their accuracy and clarity.

2. Standardized phrasing in methodologies

Certain portions of the Methods section also contain standardized language essential for scientific accuracy. For example:

- "Prioritised variants were classified based on the American College of Medical Genetics and Genomics and Association for Molecular Pathology (ACMG-AMP) guidelines for variant interpretation."

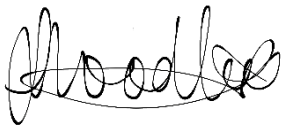
This phrasing reflects the universal guidelines used in the field of genetics and altering it would undermine the clarity and rigor expected in a scientific document.

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

It is important to emphasise that the overlapping content is either common scientific knowledge, institutional guidelines, or material that was previously written in Ms. Moodley's in the protocol and cannot be changed without losing the statements' scientific validity. These factors naturally contribute to the similarity score but do not constitute duplication or improper academic conduct.

We request that the elevated Turnitin score be reviewed in light of these considerations. If further clarification is required, we are happy to provide additional context or supporting documentation.

Sincerely,





Ms. Priya Moodley	Dr. Thandiswa Ngcungcu	02/12/2024
MSc Candidate	Supervisor	Date

Appendix F: Turnitin Report

Turnitin Test Priya Moodley - Final Write Up (MSc GM)

ORIGINALITY REPORT

21%	11%	12%	11%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Witwatersrand Student Paper	9%
2	www.mdpi.com Internet Source	1%
3	Corina M Chilibeck, Emma E Glamuzina, Casey Y-J Ung, Emma L Blakely, Robert W Taylor, Andrea L Vincent. "Albinism and a mitochondrial DNA deletion", Ophthalmic Genetics, 2020 Publication	1%
4	www.ncbi.nlm.nih.gov Internet Source	1%
5	Dejidnorov Semjid, Hyunsoo Ahn, Sapaar Bayarsaikhan, Munkhjargal Gantumur, Sanguk Kim, Jae Hoon Lee. "Identification of novel candidate genes associated with non-syndromic tooth agenesis in Mongolian families", Research Square Platform LLC, 2023 Publication	<1%

6	Submitted to Queensland University of Technology Student Paper	<1 %
7	link.springer.com Internet Source	<1 %
8	Estefanía Martínez-Barrios, Oscar Campuzano, Andrea Greco, José Cruzalegui, Georgia Sarquella-Brugada. "Cardiac channelopathies in pediatrics: a genetic update", European Journal of Pediatrics, 2024 Publication	<1 %
9	Jennifer G.R. Kromberg. "Clinical Features, Types of Albinism, and Natural History", Elsevier BV, 2018 Publication	<1 %
10	www.frontiersin.org Internet Source	<1 %
11	Submitted to Aberystwyth University Student Paper	<1 %
12	refubium.fu-berlin.de Internet Source	<1 %
13	www.science.gov Internet Source	<1 %
14	Barnell, Erica Kay. "Bioinformatic Tools to Alleviate the Annotation Bottleneck within	<1 %

Precision Oncology", Washington University in St. Louis, 2023
Publication

15	preview-cellandbioscience.biomedcentral.com Internet Source	<1%
16	wiredspace.wits.ac.za Internet Source	<1%
17	www.koreascience.or.kr Internet Source	<1%
18	"Next Generation Sequencing Based Clinical Molecular Diagnosis of Human Genetic Disorders", Springer Science and Business Media LLC, 2017 Publication	<1%
19	C. Bonnet. "Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription-PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene", Journal of Medical Genetics, 4/18/2008 Publication	<1%
20	Fierheller, Caitlin. "A Molecular Genetic Investigation of FANCI as a New Candidate Ovarian Cancer Predisposing Gene", McGill University (Canada), 2023 Publication	<1%

preview-ojrd.biomedcentral.com

21	Internet Source	<1%
22	Submitted to National School of Healthcare Science Student Paper	<1%
23	Robyn Kerr. "Identification ofP gene mutations in individuals with oculocutaneous albinism in sub-Saharan Africa", Human Mutation, 02/2000 Publication	<1%
24	www.cambridge.org Internet Source	<1%
25	etheses.bham.ac.uk Internet Source	<1%
26	www.hindawi.com Internet Source	<1%
27	www.ncbi.xyz Internet Source	<1%
28	"Oculocutaneous albinism type 2 (OCA2) with homozygous 2.7-kb deletion of the P gene and sickle cell disease in a Cameroonian family. Identification of a common TAG haplotype in the mutated P gene", Journal of Human Genetics, 09/10/2007 Publication	<1%

29	Manga, P.. "Mislocalization of Melanosomal Proteins in Melanocytes from Mice with Oculocutaneous Albinism Type 2", Experimental Eye Research, 200106 Publication	<1 %
30	www.medilib.ir Internet Source	<1 %
31	Liang, Manqi. "Identification of Novel Genes and Mechanisms Responsible for Recurrent Pregnancy Loss", McGill University (Canada), 2024 Publication	<1 %
32	Medeiros, Ana Margarida, Ana Catarina Alves, and Mafalda Bourbon. "Mutational analysis of a cohort with clinical diagnosis of familial hypercholesterolemia: considerations for genetic diagnosis improvement", Genetics in Medicine, 2015. Publication	<1 %
33	worldwidescience.org Internet Source	<1 %
34	www.grafiati.com Internet Source	<1 %
35	www.prd-journal.com Internet Source	<1 %

36	"Cassidy and Allanson's Management of Genetic Syndromes", Wiley, 2021 Publication	<1%
37	Bagheri-Fam, Stefan, Makoto Ono, Li Li, Liang Zhao, Janelle Ryan, Raymond Lai, Yukako Katsura, Fernando J. Rossello, Peter Koopman, Gerd Scherer, Oliver Bartsch, Jacob V.P. Eswarakumar, and Vincent R. Harley. "FGFR2 mutation in 46,XY sex reversal with craniosynostosis", Human Molecular Genetics, 2015. Publication	<1%
38	Chaoyi Li, Qian Chen, Junjiao Wu, Jie Ren, Mengfei Zhang, Huakun Wang, Jinchen Li, Yu Tang. "Identification and characterization of two novel non-coding Tyrosinase (TYR) gene variants leading to Oculocutaneous albinism type 1", Journal of Biological Chemistry, 2022 Publication	<1%
39	Fernandes, Ana Rita Albuquerque. "Development, Validation and Implementation of an NGS Gene Panel Approach for the Diagnosis of Renal Amyloidoses", Universidade de Tras-os-Montes e Alto Douro (Portugal), 2024 Publication	<1%
40	Junyu Zhang, Yanyi Yao, Haixian He, Jun Shen. "Clinical Interpretation of Sequence Variants",	<1%

	Current Protocols in Human Genetics, 2020 Publication	
41	Sanna Gudmundsson, Moriel Singer-Berk, Nicholas A. Watts, William Phu et al. "Variant interpretation using population databases: Lessons from gnomAD", Human Mutation, 2021 Publication	<1 %
42	Tuli, Abbas M., Robert K. Valenzuela, Erasmus Kamugisha, and Murray H. Brilliant. "Albinism and disease causing pathogens in Tanzania: Are alleles that are associated with OCA2 being maintained by balancing selection?", Medical Hypotheses, 2012. Publication	<1 %
43	d-nb.info Internet Source	<1 %
44	diseasewiki.hupo.org.cn Internet Source	<1 %
45	downloads.hindawi.com Internet Source	<1 %
46	epdf.tips Internet Source	<1 %
47	orbi.uliege.be Internet Source	<1 %
48	www.ajol.info Internet Source	

		<1%
49	www.escholar.manchester.ac.uk Internet Source	<1%
50	www.life-science-alliance.org Internet Source	<1%
51	www.researchgate.net Internet Source	<1%
52	www.researchsquare.com Internet Source	<1%

Exclude quotes	On	Exclude matches	< 8 words
Exclude bibliography	On		

Appendix G: Frontiers in Genetics Author guidelines

Author guidelines

Writing and formatting

Title

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

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

Authors should avoid:

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

Abstract

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

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

Manuscript length

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

Sections

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

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

Introduction
Succinct, with no subheadings.

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

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

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

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

Language editing

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

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

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

Editage
We recommend the language-editing service provided by our external partner Editage. These

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

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

Frontiers**推荐您使用在英**语语言编辑和校对领域具有悠久历史和良好口碑的查尔斯沃思作者服务。此项服务由第三方为您提供，Frontiers**中国作者通**过此链接提交稿件时可获得**10%的特**别优惠: **www.cwauthors.com.cn/frontiers**.

Language style

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

Inclusive language guidelines

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

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

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

In general, seek to avoid

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

- any form of language that suggests a particular culture or group as the default or standard.

And where feasible:

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

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

Guidelines for artificial intelligence and related technologies

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

Use of AI to create written or visual content

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

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

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

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

Abbreviations and nomenclatures

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

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

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

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

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

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

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

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

Enhancing search engine optimization (SEO)

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

- Include a few of your article's keywords in the title of the article
- Do not use long article titles
- Pick 5-8 keywords using a mix of generic and more specific terms on the article subject(s)
- Use the maximum amount of keywords in the first two sentences of the abstract
- Use some of the keywords in level 1 headings

References

Preparing and formatting references

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

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

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

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

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

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

Harvard reference style (author-date)

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

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

EXAMPLES:

- Article in a print journal:

Sondheimer, N., and Lindquist, S. (2000). Rnq1: an epigenetic modifier of protein function in yeast. Mol. Cell. 5, 163-172.

(Sondheimer and Lindquist, 2000)

- Article in an online journal:

Tahimic, C.G.T., Wang, Y., Bikle, D.D. (2013). Anabolic effects of IGF-1 signaling on the skeleton. Front. Endocrinol. 4:6. doi: 10.3389/fendo.2013.00006

(Tahimic et al., 2013)

- Article or chapter in a book:

Sorenson, P. W., and Caprio, J. C. (1998). "Chemoreception," in The Physiology of Fishes, ed. D. H. Evans (Boca Raton, FL: CRC Press), 375-405.

(Sorenson and Caprio, 1998)

- Book:

Cowan, W. M., Jessell, T. M., and Zipursky, S. L. (1997). Molecular and Cellular Approaches to Neural Development. New York: Oxford University Press.

(Cowan et al., 1997)

- Abstract:

Hendricks, J., Applebaum, R., and Kunkel, S. (2010). A world apart? Bridging the gap between theory and applied social gerontology. Gerontologist 50, 284-293. Abstract retrieved from Abstracts in Social Gerontology database. (Accession No. 50360869)

(Hendricks et al., 2010)

- Website:

World Health Organization. (2018). E. coli. https://www.who.int/news-room/fact-sheets/detail/e-coli [Accessed March 15, 2018].

(World Health Organisation, 2018)

- Patent:

Marshall, S. P. (2000). Method and apparatus for eye tracking and monitoring pupil dilation to evaluate cognitive activity. U.S. Patent No 6,090,051. Washington, DC: U.S. Patent and Trademark Office.

(Marshall, 2000)

- Data:

Perdiguero P, Venturas M, Cervera MT, Gil L, Collada C. Data from: Massive sequencing of Ulms minor's transcriptome provides new molecular tools for a genus under the constant threat of Dutch elm disease. Dryad Digital Repository. (2015) http://dx.doi.org/10.5061/dryad.ps837

(Perdiguero et al., 2015)

- Theses and dissertations:

Smith, J. (2008) Post-structuralist discourse relative to phenomological pursuits in the deconstructivist arena. [dissertation/master's thesis]. [Chicago (IL)]: University of Chicago

(Smith, 2008)

- Preprint:

Smith, J. (2008). Title of the document. Preprint repository name [Preprint]. Available at: https://persistent-url (Accessed March 15, 2018).

(Smith, 2008)

- One author:

Hesse-Biber, S. (2010). Qualitative Approaches to Mixed Methods Practice. Qualitative Inquiry, 16(6), 455-468. https://doi.org/10.1177/1077800410364611

(Hesse-Biber, 2010)

- Two authors:

Duvail, S., Hamerlynck, O. (2007) The Rufiji River flood: plague or blessing?. Int. J. Biometeorol. 52, 33–42. https://doi.org/10.1007/s00484-007-0105-8

(Duvail and Hamerlynck, 2007)

- More than two authors:

Leemhuis C, Thonfeld F, Näschen K, Steinbach S, Muro J, Strauch A, López A, Daconto G, Games I, Diekkrüger B. (2017) Sustainability in the Food-Water-Ecosystem Nexus: The Role of Land Use and Land Cover Change for Water Resources and Ecosystems in the Kilombero Wetland, Tanzania. Sustainability. 9(9):1513. https://doi.org/10.3390/su9091513

(Leemhuis et al., 2017)

- Same author(s); same year:

Huang, X-G. (2016a). Electromagnetic fields and anomalous transports in heavy-ion collisions—a pedagogical review. Rep. Prog. Phys. 79 076302. doi: 10.1088/0034-4885/79/7/076302 Huang, X-G. (2016b). Simulating Chiral Magnetic and Separation Effects with Spin-Orbit Coupled Atomic Gases. Scientific Reports. 6:20601. doi: 10.1038/srep20601

(Huang, 2016a, 2016b)

- Same author(s); different years:

Sedrakian, A. (2007). The physics of dense hadronic matter and compact stars. Progress in Particle and Nuclear Physics. 58(1):168-246. doi: 10.1016/j.ppnp.2006.02.002

Sedrakian, A. (2016). Axion cooling of neutron stars. Phys. Rev. D 93:065044. doi. 10.1103/PhysRevD.93.065044

(Sedrakian, 2007, 2016)

- Same first author; different author list:

Quimque, M. T., Notarte, K. I., Letada, A., Fernandez, R. A., and Pilapil, D. Y. 4th., Pueblos, K.R., Agbay, J.C., Dahse, H.M., Wenzel-Storjohann, A., Tasdemir, D., Khan, A., Wei, D.Q., Gose Macabeo, A.P. (2021a). Potential Cancer- and Alzheimer's Disease-Targeting Phosphodiesterase Inhibitors fromUvaria alba: Insights from In Vitro and Consensus Virtual Screening. ACS Omega. 6, 8403–8417. doi: 10.1021/acsomega.1c00137

Quimque, M. T. J., Notarte, K. I. R., Fernandez, R. A. T.,Mendoza,M. A. O., Liman, R. A. D., Lim, J. A. K., et al. (2021b). Virtual screening-driven drug discovery of SARSCoV2 enzyme inhibitors targeting viral attachment, replication, post-translational modification and host immunity evasion infection mechanisms. J Biomol Struct Dyn. 39, 4316–4333. doi: 10.1080/07391102.2020.1776639

(Quimque et al., 2021a; Quimque et al., 2021b)

- Different authors; same surname:

Khan, S. M., Khan, M., Alouffi, A., Almutairi, M. M., Numan, M., Ullah, S., et al. (2023). Phylogenetic position of Haemaphysalis kashmirensis and Haemaphysalis cornupunctata, with Notes on Rickettsia spp. Genes. 14, 360. doi: 10.3390/genes14020360 Khan, Z., Shehla, S., Alouffi, A., Kashif Obaid, M., Zeb Khan, A., Almutairi, M. M., et al. (2022). Molecular survey and genetic characterization of Anaplasma marginale in ticks collected from livestock hosts in Pakistan. Animals. 12, 1708. doi: 10.3390/ani12131708

(Khan Z. et al., 2022; Khan S. M. et al., 2023)

- Publishing in a Humanities and Social Sciences journal:

Farrell, H. (2012). The Consequences of the Internet for Politics. Annual Review Political Science. 15, 35-52. http://dx.doi.org/10.1146/annurev-polisci-030810-110815

(Farrell, 2012, p. 40)

- Personal communications:

“We thank L. Li (personal communication, December, 2018) for noting this ambiguity.”

Citation of retracted articles

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

Guidelines for citing retracted articles:

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

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

Acknowledgements

This is a short text to acknowledge the contributions of specific colleagues, institutions, or agencies that aided the efforts of the authors. Should the content of the manuscript have

previously appeared online, such as in a thesis or preprint, this should be mentioned here, in addition to listing the source within the reference list.

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

Templates

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

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

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

Figures, tables, and images

Figures, tables, and images: rights and permissions

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

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

Figures and images: style guidelines

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

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

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

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

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

Captions

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

Image size and resolution requirements

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

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

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

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

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

Format and color image mode

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

Images of chemical structures

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

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

Table requirements and style guidelines

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

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

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

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

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

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

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

Accessibility

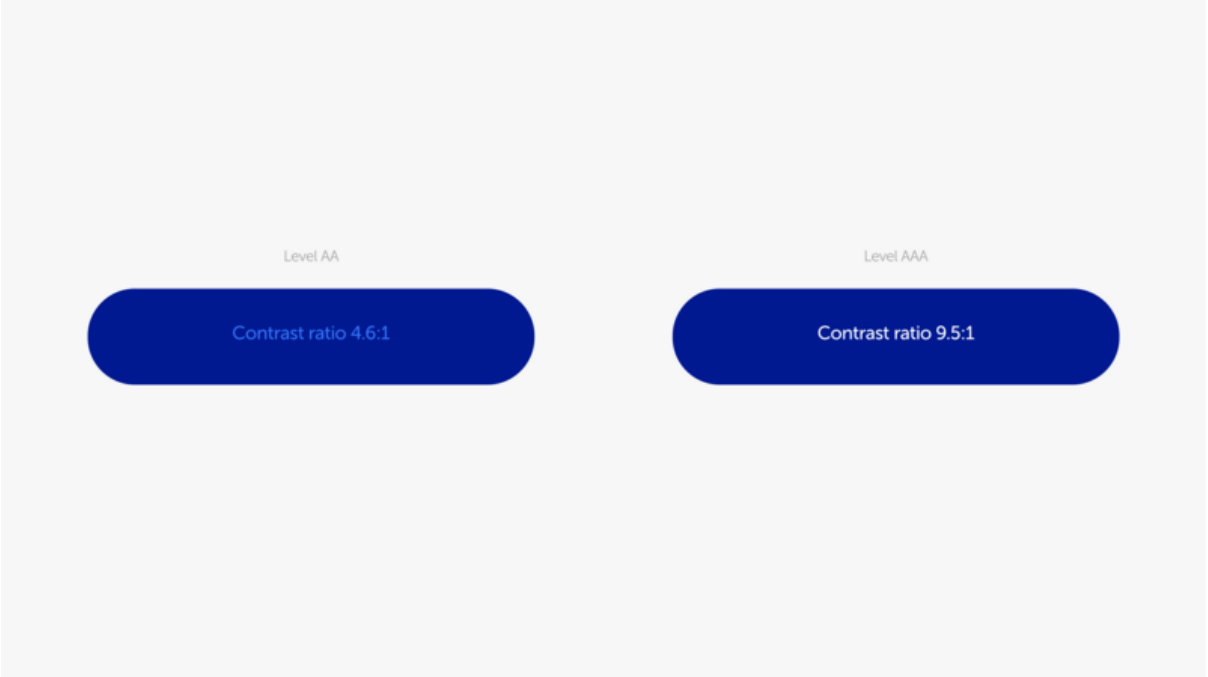
We encourage authors to make the figures and visual elements of their articles accessible for the visually impaired. Effective use of color can help people with low visual acuity, or color blindness, understand all the content of an article.

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

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

WC3 recommends the following contrast ratio levels:

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

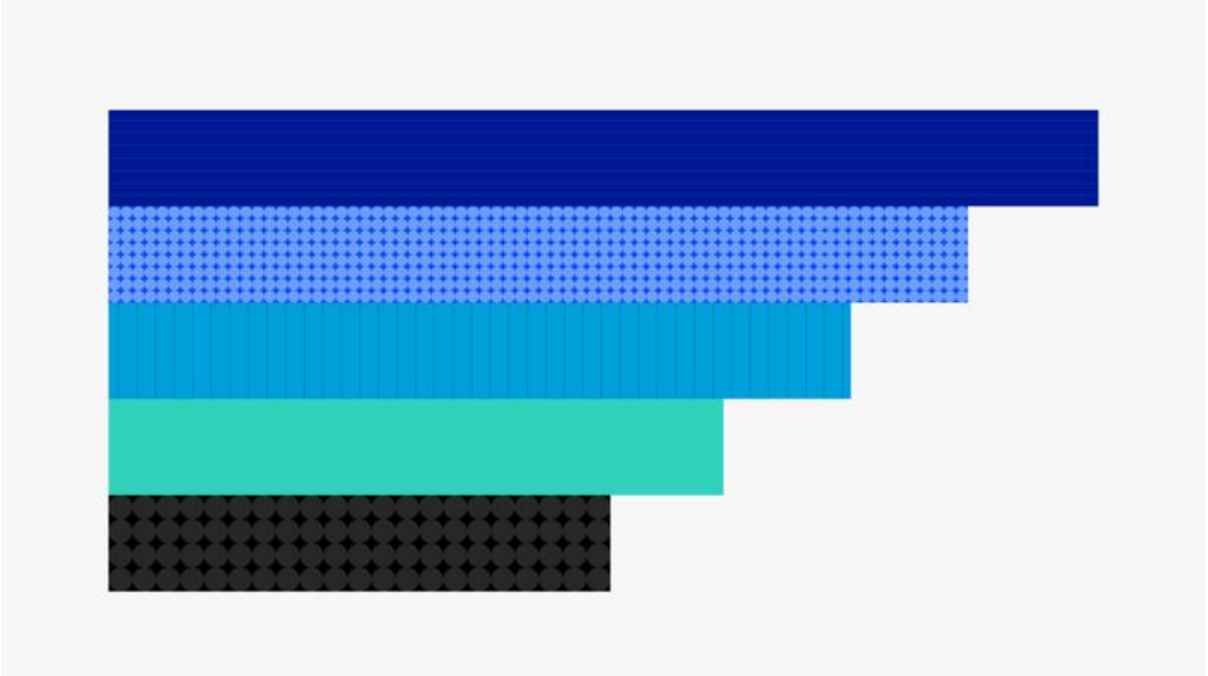


You can verify the contrast ratio of your palette with these online ratio checkers:

- WebAIM
- Color Safe

Avoid using red or green indicators
More than 99% of color-blind people have a red-green color vision deficiency.

Avoid using only color to communicate information
Elements with complex information like charts and graphs can be hard to read when only color is used to distinguish the data. Try to use other visual aspects to communicate information, such as shape, labels, and size. Incorporating patterns into the shape fills also make differences clearer; for an example please see below:



Supplementary material

Selecting supplementary material

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

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

File type requirements

The supplementary material can be uploaded as:

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

Technical requirements for supplementary images:

- 300 DPIs
- RGB color mode.

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

Submitting information

Choosing where to submit

Open access and copyright

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

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

Registration with Frontiers

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

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

Article type

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

Keywords

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

CrossMark policy

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

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

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

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

Authors and affiliations

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

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

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

Correspondence

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

Example: Max Maximus*
maximus@iuscience.edu

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

Equal contributions

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

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

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

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

Consortium/group and collaborative authors

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

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

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

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

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

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

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

Scope statement

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

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

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

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