

INVESTIGATING THE GENETIC CAUSE OF OCULOCUTANEOUS ALBINISM TYPE 2 IN INDIVIDUALS OF AFRICAN DESCENT THROUGH EXOME SEQUENCING

Phophi M. Mmbi

Student number: 1647122

Supervisor: Dr Thandiswa G. Ngcungcu

Co-Supervisors: Dr Robyn A. 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, 2024

NATIONAL HEALTH LABORATORY SERVICE

School of Pathology, University of the Witwatersrand



DIVISION OF HUMAN GENETICS



UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Hospital Street, Johannesburg, 2001 | PO Box 1038, Johannesburg, 2000
[T] +27 11 489 9223 | [M] +27 78 080 8841 | [F] +27 11 489 9226 | [E] human.genetics@nhls.ac.za

Dear examiner,

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

Thank you for agreeing to examine the research report titled "Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of African descent through exome sequencing".

Ms. Phophi Mmbi, a candidate for MSc (Med) in Genomic Medicine, is submitting her Research Report in a "Submissible" format. The Research Report contributes 33% towards the final mark in this MSc degree requirements. The Research Protocol was assessed and approved by the Faculty of Health Science and is attached below (Appendix D) for a detailed literature review, but is not for examination purposes.

The candidate and her supervisors, Dr. Thandiswa Ngcungcu, Dr. Phelelani Mpangase and Dr. Robyn Kerr, have chosen to submit the Research Report to the journal Frontiers in Genetics, therefore the manuscript was formatted according to the author guidelines of this journal. These guidelines have been attached for your reference (Appendix F). There are deviations purposively made from the journal submission guidelines (e.g. the tables and figures are placed in the text and not submitted separately as per the Frontiers guidelines) for ease of reading. Additionally, line numbers are not inserted and a line spacing of 1.5 was used instead of a line spacing of 1, also to enhance readability. Supporting documents are called Appendices throughout the manuscript rather than supplementary documents.

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

Yours sincerely

Student

Supervisor 1

Supervisor 2

Supervisor 3

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, Phophi Mmbi, 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.

A handwritten signature in black ink, appearing to read 'Phophi Mmbi', is written over a horizontal line.

Phophi Mmbi

Student number: 1647122

Contribution of the candidate to the paper

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

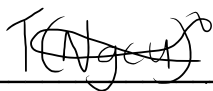
I, Phophi Mmbi, student number 1647122, declare that this Research Report, entitled "Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of African descent through exome sequencing", is my own work and that I contributed significantly towards the research findings presented in the paper intended for publication below.



Signature of Student

26/02/2024

Date



Signature of Primary Supervisor

26/02/2024

Date

Agreement by co-authors:

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

Authors	Name	Signature	Date
1st	Ms Phophi Mmbi		26/02/2024
2nd	Dr Thandiswa Ngcungcu		26/02/2024
3rd	Dr Robyn Kerr		26 February 2024
4th	Dr Phelelani Mpangase		26 February 2024

DEDICATION

To my mother, Stellah Mmbi.

ABSTRACT

Oculocutaneous Albinism type 2 (OCA2) is a hypopigmentation disorder caused by variants in the *OCA2* gene. A 2.7kb intragenic deletion is known to be the common variant that causes OCA2 in individuals of African descent. This variant accounts for 78% of OCA2-causing variant in the Southern African population. The diagnostic utility for individuals who tested either negative or heterozygous for the common 2.7kb deletion remained unsolved. This study reports on the identification of the NM_000275.3:c.1503+5G>A variant found in 3/8 (37.5%) non-2.7kb deletion OCA2 chromosomes within a small sample of OCA2 patients from the Southern African population, strongly suggesting the possibility of a second common variant, pending a larger screening study. This variant has been previously described in this population, but new variant interpretation tools have now permitted its reclassification. Furthermore, limitations encountered during the interpretation of data step in this study highlight the importance of informative clinical and phenotypic data for improved and sensible interpretation of genetic results.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my esteemed supervisors, Dr Thandiswa Ngcungcu, Dr Phelelani Mpangase, and Dr Robyn Kerr, whose steadfast support, guidance, and expertise have been instrumental in the successful completion of this thesis. Your commitment to excellence, insightful feedback, and encouragement have truly shaped this research journey and enriched my academic experience. To Dr Thandiswa Ngcungcu, thank you for the hands-on experience you allowed me to have with this project, your guidance through all the laboratory work is truly appreciated. To Dr Phelelani Mpangase, thank you for introducing me to the world of bioinformatic analysis, I am grateful for the incredible opportunity and your patience through it all. To Dr Robyn Kerr, thank you for sharing your wealth of knowledge on Albinism with me. Your exceptional guidance and support in refining my writing skills has been invaluable to me. Thank you all for being an integral part of this journey, and for helping turn the pages of this significant chapter in my academic life.

My gratitude extends to the University of the Witwatersrand for providing the Postgraduate Merit Award funding opportunity that made it possible for me to undertake my studies. And a special thanks again to Dr Thandiswa Ngcungcu for funding this project through your MRC SIR grant.

A special note of appreciation goes to Dr Maria Mudau and Dr Lindie Lamola for your assistance during this year. You both expressed an unwavering willingness to assist me when needed. Your guidance and generosity in sharing your knowledge have been proven invaluable to me. I am sincerely grateful to both of you for contributing to this transformative experience.

Lastly, I am deeply thankful to my family and close friends for their unwavering support, understanding, and encouragement. Your patience during challenging times and celebration during the triumphs have been the pillars of my strength. Your belief in me has been a driving force, motivating me to persevere through the highs and lows of this academic journey.

TABLE OF CONTENTS

Declaration	i
Contribution of the Candidate to the Paper	ii
Dedication	iii
Abstract	iv
Acknowledgments	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
List of Abbreviations	x
1 Introduction	2
2 Methods and Materials	3
2.1 Sample Selection	3
2.2 Genotyping the 2.7kb Deletion	4
2.3 Whole Exome Sequencing	4
2.4 Variant Calling, Prioritisation and Classification	5
3 Results	5
3.1 Genotyping the common 2.7kb deletion	5
3.2 Whole Exome Sequencing and variant analysis	5
3.3 Other pigment genes	10
3.4 Overall result summation	12
4 Discussion	12
4.1 Variants in <i>OCA2</i>	12
4.2 Variants in <i>TYRP1</i>	14
4.3 The remaining unresolved cases	15
4.4 Limitations and future directions	15
5 Conclusion	16
Data Availability Statement	16
Ethics Statement	16
Author Contributions	17
Funding	17
Acknowledgments	17
Conflict Of Interest	17
References	17

6	Appendices	22
	Appendix A: Supplementary Material	22
	Appendix B: HREC (Medical), University of the Witwatersrand (M230778)	24
	Appendix C: HREC (Medical), University of the Witwatersrand (M210203)	26
	Appendix D: Approved Research Protocol	28
	Appendix E: Turnitin Report and Plagiarism Declaration	46
	Appendix F:Frontiers in Genetics Author guidelines	54

LIST OF FIGURES

1	Genotyping of the common 2.7kb deletion in the <i>OCA2</i> gene	6
2	Variants called within the whole exome in a cohort of eight African <i>OCA2</i> individuals	6
3	Types of variants called within the <i>OCA2</i> gene from whole exome sequencing data	7
4	IGV visualisation of NM_000275.3:c.1503+5G>A	9
5	Sanger sequencing validation for NM_000275.3:c.1503+5G>A	10
S1	IGV visualization of the 2.7kb deletion region	23

LIST OF TABLES

1	Patient clinical and family history	4
2	<i>OCA2</i> (NM_000275.3) variants prioritised using ACMG/AMP guidelines	8
3	<i>TYRP1</i> (NM_000550.3) variants prioritised using ACMG/AMP guidelines	11
4	Summary of results and variant profile per case	13
S1	PCR primer sequences designed to amplify the NM_000275.3:c.1503+5G>A variant region	22

LIST OF ABBREVIATIONS

Abbreviation	Full name
Array-CGH	Comparative Genome Hybridization Microarray
BOCA	Brown Oculocutaneous Albinism
<i>HERC2</i>	Hect Domain And Rcc1-Like Domain 2 gene
INDEL	Insertion-Deletion
MAF	Minor Allele Frequency
<i>MC1R</i>	Melanocortin 1 Receptor gene
NGS	Next Generation Sequencing
NICU	Neonatal Intensive Care Unit
OCA	Oculocutaneous Albinism
<i>OCA2</i>	Oculocutaneous Albinism Type 2 gene
OCA2	Oculocutaneous Albinism Type 2
OCA3	Oculocutaneous Albinism Type 3
PCR	Polymerase Chain Reaction
ROCA	Rufous Oculocutaneous Albinism
SNV	Single Nucleotide Variant
<i>TYR</i>	Tyrosinase gene
<i>TYRP1</i>	Tyrosinase-Related Protein 1 gene
VCF	Variant Call File
VEP	Variant Effect Predictor
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing

Research report in the format of a “submissible” paper

To be submitted to Frontiers in Genetics Journal

Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of African descent through exome sequencing

Phophi M. Mmbi^{1,*}, Robyn A. Kerr¹, Phelelani T. Mpangase² 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, University of the Witwatersrand, Johannesburg, 2193, South Africa*

Correspondence*:

Phophi M. Mmbi

1647122@students.wits.ac.za

Thandiswa G. Ngcungcu

Thandiswa.Ngcungcu@wits.ac.za

ABSTRACT

Background: Oculocutaneous Albinism type 2 (OCA2) is a hypopigmentation disorder caused by variants in the *OCA2* gene. A 2.7kb deletion has been shown to be the common variant that causes OCA2 in individuals of African descent. This variant accounts for 78% of OCA2-causing variants in the Southern African population. For individuals who test either negative or heterozygous for the common 2.7kb deletion, no further molecular testing could be routinely offered in our resource-constrained Diagnostic Laboratory and the unknown variant/s remain unidentified. This research study aimed to investigate genes associated with skin pigmentation, primarily the *OCA2* gene, to identify other possible variants that cause or contribute to the OCA2 phenotype within the South African context.

Methods: Genotyping of the 2.7kb deletion was carried out, followed by whole exome sequencing and screening of *OCA2*, *HERC2*, *TYR*, and *TYRP1* genes. Bioinformatic analyses were performed to prioritise potential pathogenic variants. Pathogenicity was determined using ACMG guidelines. Sanger sequencing was used to validate the clinically significant variant identified within the *OCA2* gene.

Results: One likely pathogenic variant, NM_000275.3:c.1503+5G>A, was identified in the *OCA2* gene and has a high frequency in this small cohort (3/7 chromosomes). This variant was previously described in African *OCA2* individuals and classified as a benign variant at the time, but current variant interpretation tools have now allowed for its reclassification as a disease-causing variant. Two common rufous oculocutaneous albinism (ROCA) causing variants, NM_000550.3:c.497C>G (p.Ser166Ter) (rs10489413) and NM_000550.3:c.1103del (p.Lys368fs), were identified in the *TYRP1* gene in one patient.

Conclusion: This study reports on the identification of the NM_000275.3:c.1503+5G>A variant within a Southern African *OCA2* group of patients and suggests this is a second common variant in this population, pending a larger screening study. Furthermore, this study illustrates the utility of newer variant interpretation tools which can allow for reclassification of variants, and also highlights the importance of good clinical/phenotype data at the time of test requesting as this data can often be used to inform accurate, targeted molecular testing.

Keywords: Oculocutaneous Albinism Type 2, *OCA2*, African, Exome sequencing, Common variant

1 INTRODUCTION

Albinism refers to a group of inherited hypopigmentation disorders that can arise from syndromic or non-syndromic gene-disease associations. There are two types of non-syndromic albinism: (i) ocular albinism, a rarer form mainly affecting the eyes; and (ii) oculocutaneous albinism (OCA), a common form that affects the eyes, hair, and skin. OCA shows locus heterogeneity and presents as eight clinical types (OCA1-OCA8) differentiated by phenotype and genetic causes. (Grønskov et al., 2007; Kromberg et al., 2012; Montoliu et al., 2014; Thomas et al., 2023). The most common OCA types observed in African populations are *OCA2* and *OCA3* (Rufous OCA - ROCA) (Kromberg et al., 2012; Manga et al., 2001, 1997; Stevens et al., 1997). *OCA2* is caused by variants in the *OCA2* gene (OMIM 611409), which encodes for a transmembrane protein, the P protein, involved in the processing and regulation of melanosomal proteins (Chen et al., 2002; Toyofuku et al., 2002). *OCA3* is caused by variants in the *TYRP1* gene (OMIM 115501) (Manga et al., 1997).

Phenotypic differences between the two types are distinctive, with OCA3 (ROCA) exhibiting more pigmentation than OCA2 (Kromberg et al., 2012; Manga et al., 1997). The clinical features, such as visual impairment and predisposition to skin cancer, are more likely to manifest in individuals with OCA2 as compared to OCA3 (Kromberg et al., 2008, 2012).

The incidence of OCA2 in black South Africans is 1/3900 births, but the prevalence varies across Southern Africa (Kromberg and Jenkins, 1982). The most common cause of OCA2 in sub-Saharan Africans is a 2.7kb deletion mapped to chr15:28262863-28265786 (GRCh37), which includes the deletion of exon 7 of the *OCA2* gene (Durham-Pierre et al., 1994; Stevens et al., 1995; Morice-Picard et al., 2014). This deletion accounts for 78% of mutated OCA2 chromosomes and has a carrier frequency of 1 in 40 within the black Southern Africans population (Stevens et al., 1995, 1997). A large proportion of disease-causing OCA2 variants remain to be identified (Kerr et al., 2000; Spritz et al., 1995). To our knowledge, no recent studies have been conducted to discover other variants that cause or contribute to the OCA2 phenotype in the Southern African population.

Limitations to variant discovery in earlier studies have been attributed to the methodology, detection accuracy of genetic testing approaches used, and possible influence of modifier genes such as *HERC2* and *MC1R* (Kerr et al., 2000; King et al., 2003; Oetting et al., 2005; Spritz et al., 1995). Recent technological advancements have made it possible to analyse multiple genes and adjacent regions simultaneously through next-generation sequencing (NGS) techniques and bioinformatic tools. This study therefore aimed to investigate pigment genes (particularly the *OCA2* gene) using whole exome sequencing (WES) to identify causative variants in individuals who tested either heterozygous or negative for the common 2.7kb deletion.

2 METHODS AND MATERIALS

2.1 Sample Selection

A total of eight (P1-P8) de-identified, residual DNA samples from individuals who tested heterozygous (n=7) or negative (n=1) for the 2.7kb deletion in the *OCA2* gene were selected for this study. All participants were unrelated individuals of African descent who were referred for routine OCA2 diagnostic testing at the Division of Human Genetics, National Health Laboratory Service, Braamfontein. Although minimal information was sent with each test request, each patient's family history and/or phenotype data were documented as per the available patient file. Patient files for P1, P2, and P8 could unfortunately not be located from within the Division's filing system (Table 1).

Table 1: Patient clinical and family history

Patient	Family history	Clinical/phenotype data*
P1	-	-
P2	-	-
P3	-	"presents with albinism"
P4	"affected mother; unaffected father"	partial albinism; no nystagmus; "medium fair complexion; blackish brown hair"
P5	"unaffected parents, one affected sister, and six unaffected siblings"	"presents with albinism"
P6	-	"?albinism"
P7	Family history of albinism	"NICU baby; clinically pale hair and eye colour"
P8	-	-

* as per the testing requisition slip sent from the referring clinician

2.2 Genotyping the 2.7kb Deletion

A polymerase chain reaction (PCR) to confirm the presence or absence of the 2.7kb deletion in the *OCA2* gene was performed as described in earlier studies (Durham-Pierre et al., 1994; Puri et al., 1997). The eight samples were run together with two controls (a heterozygote for the 2.7kb deletion and a sample that is negative for the deletion) and a no template control.

2.3 Whole Exome Sequencing

Sample quantity and quality were determined using Nanodrop and Qubit™ 3.0 Fluorometer. Libraries were generated using the Ion AmpliSeq™ Exome RDY Kit as per manufacturer protocol (Thermo Fisher Scientific, MAN0010084). In this process, the DNA samples (100ng) were amplified, digested, and ligated to IonCode™ Barcode Adapters to form barcoded libraries. The barcoded library was then purified using AMPure XP Bead Reagent (Beckman Coulter, CA, USA). Purified barcoded libraries were amplified, and then quantified using the Tape station with High Sensitivity D5000 Reagents (Agilent Technologies). The libraries were diluted to 22pM and pooled together in pairs for loading onto an Ion 540™ Chip. Libraries were templated using the Ion 540™ Kit-Chef following the manufacturer's instructions (MAN0010851, Thermo Fisher Scientific, South Africa) and sequenced with the Ion GeneStudio™ S5 Prime System (Thermo Fisher Scientific, South Africa). The Ion GeneStudio™ S5 Prime System carries out a comprehensive internal WES analysis, from sequencing, quality control, alignment (to GRCh37) and variant calling.

2.4 Variant Calling, Prioritisation and Classification

2.4.1 Single nucleotide variant analysis

The variant call file (VCF) for each sample, obtained from the Ion GeneStudio™ S5 Prime System, was annotated using a local installation of the Ensembl Variant Effect Predictor (VEP, version 110) and the GRCh37 reference genome (McLaren et al., 2016). Annotated variants were filtered and prioritised following a tiered approach based on: (i) the gene (initially *OCA2*, then other genes were analysed where no variant that met the filtering criteria was identified in *OCA2*); (ii) predicted consequence to resultant protein (such as frameshift, missense, and splice site variants); and (iii) a minor allele frequency of $\pm 5\%$ using the 1000 Genomes Project and GnomAD databases. Prioritised variants were classified based on the American College of Medical Genetics and Genomics (ACMG) guidelines for variant interrogation (Abou Tayoun et al., 2018; Brnich et al., 2020; Richards et al., 2015). Quality assessment using the Integrative Genomics Viewer (IGV v2.16.2) (Robinson et al., 2017) was conducted for each prioritised variant to further assess the depth of coverage, confirm zygosity, and exclude artefacts. Analysis of other pigment genes followed the same filtering criteria applied to *OCA2*. Genes chosen for analysis include *HERC2*, *TYR*, and *TYRP1*. Variants with clinically significant classifications were validated through Sanger sequencing.

2.4.2 Copy number variant analysis

Copy number variant (CNV) analysis was carried out using the Ion Reporter™ software (v5.20). This analysis software reportedly picks up copy number variants at a minimum coverage of 0.01x (MAN0028321 A.0, Thermo Fisher Scientific).

3 RESULTS

3.1 Genotyping the common 2.7kb deletion

The presence of the 2.7kb deletion was confirmed in 7/8 samples included in this study (Figure 1) in agreement with the molecular testing data available from the Diagnostics Laboratory. All individuals, except P7 were heterozygous for the deletion. Individual P7 was negative for the deletion.

3.2 Whole Exome Sequencing and variant analysis

The mean depth coverage obtained from the WES was within the range of 44 to 145X, with a mean read length of 209 bp. For all the samples, more than 99% of the sequenced reads

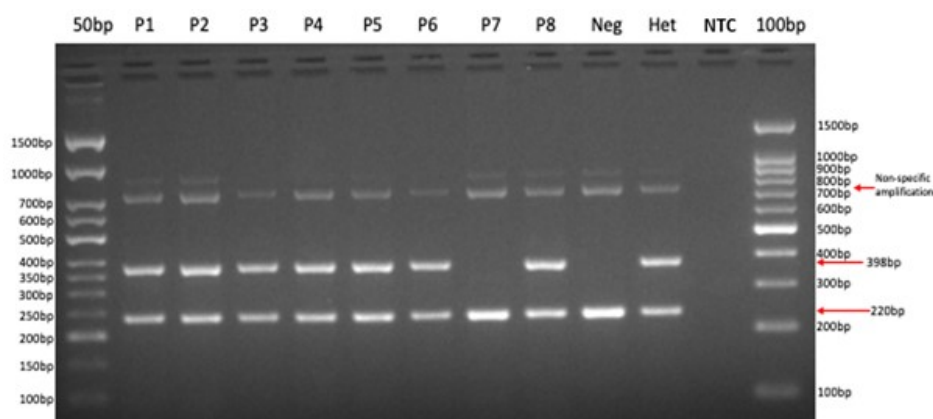


Figure 1: Genotyping of the common 2.7kb deletion in the *OCA2* gene. PCR results were visualised on a 3% agarose gel, with bands of 398 bp (2.7kb deletion) and 220 bp (normal allele) in size. Negative (Neg) and heterozygous (Het) controls for the 2.7kb deletion were included as allelic references. P1-6 and P8 were heterozygous whilst P7 was negative for the 2.7kb deletion. Bands seen in the 700 - 1000 bp range represent non-specific amplification, but do not interfere with result interpretation. A no template control (NTC) was included in the PCR set up; there was no contamination identified within the PCR.

were aligned to the GRCh37 reference sequence and less than 1% were not aligned. Over 400,000 variants were identified with at least 250,000 variants called per individual sample from the resulting VCF files. The majority of the variants were missense, synonymous and intronic in nature, with the highest class being SNVs (Figure 2 A, B).

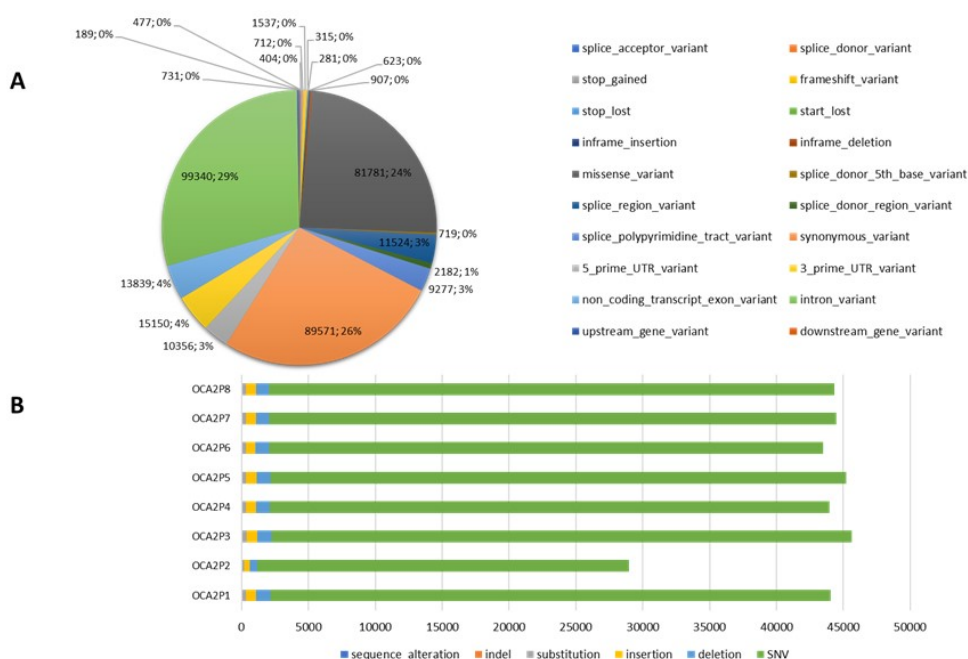


Figure 2: Variants called within the whole exome in a cohort of eight African *OCA2* individuals. **A.** Overall variant consequences within the whole exome annotated using VEP. Variant types were collated for each sample and the sum presence of each variant type was visualised. **B.** Summarised variant classes for whole exome per individual. The variant classes identified most frequently per sample were SNVs, followed by deletions and insertions.

A total of 66 variants were found in the *OCA2* gene across the eight individuals (22 unique variants, once duplicates were removed) (Figure 3). Following this, variants with a potentially unharmed consequence (intronic and synonymous) were removed, leaving 18 variants across two consequent types (missense and splicing variants). After accounting for duplicates, only five SNVs spread across the eight patients were detected within the *OCA2* gene (Table 2). No CNVs (including the common 2.7kb deletion) or indels were identified in the *OCA2* gene. The 2.7kb deletion region (chr15:28262863-28265786) was examined using the IGV with P3 (heterozygous for the deletion) and P7 (negative for the deletion). A decrease in coverage was observed in P3, but it was not statistically significant (Figure S1).

Finally, variants were filtered based on MAF. Two of these variants NM_000275.3:c.1183-4A>G (rs10852218) and NM_000275.3:c.913C>T (p.Arg305Trp) (rs1800401) were notably common in population databases (1000 Genome Project and GnomAD). NM_000275.3:c.1183-4A>G had a frequency of 0.42 and 0.47; whereas NM_000275.3:c.913C>T was reported as 0.13 and 0.14 respectively within the African population (1000 Genome Project_AFR and GnomAD_AFR). These variants were subsequently excluded from further analysis. The remaining three variants, NM_000275.3:c.1007C>T (p.Ala336Val) (rs34010619) (n=1), NM_000275.3:c.2165T>C (p.Ile722Thr) (rs1800417) (n=1), and NM_000275.3:c.1503+5G>A (rs368124046) (n=3) (Table 2), all with a MAF frequency ≤ 0.05 , were prioritised for classification using the ACMG guidelines. PM2 was thus applied for all three variants for low allele frequencies within the African population. In all individuals, these variants were heterozygous.

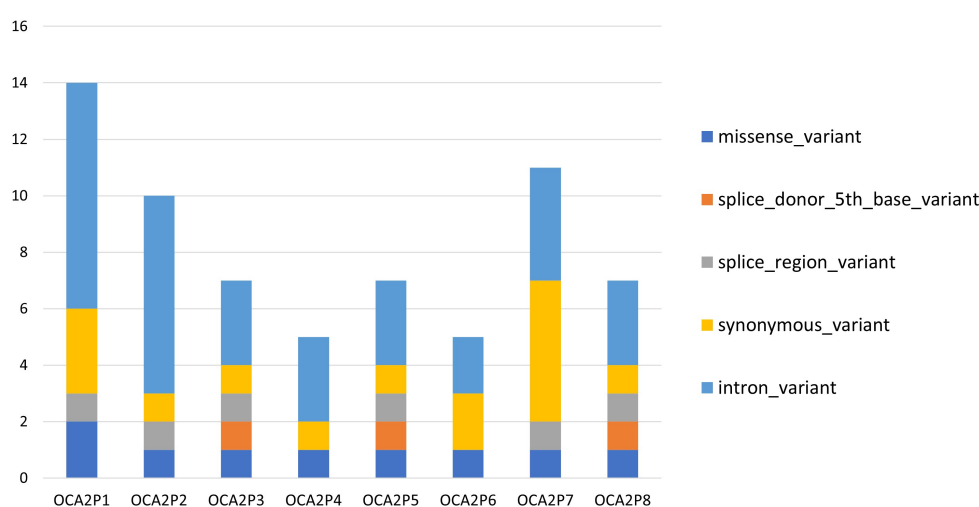


Figure 3: Types of variants called within the *OCA2* gene from whole exome sequencing data. A total of 66 variants across eight samples were identified and categorised by variant type and consequence.

Table 2: *OCA2* (NM_000275.3) variants prioritised using ACMG/AMP guidelines

Gene	Variation		SNP	Allele frequency ^a	Occurrence ^b	Consequence	In silico prediction	Classification	Evidence	Ref
	HGVSc	HGVSp								
<i>OCA2</i>	C.1503+5G>A	-	rs368124046	<0	3/8	Splice donor	0.76 ^c	Likely Pathogenic	PS3, PM2 PP3, PP5	1
<i>OCA2</i>	C.1007C>T	p.Ala336Val	rs34010619	0.03	1/8	Missense	0.19 ^d	Likely Benign	PM2, BP4 BP6	2
<i>OCA2</i>	C.2165T>C	p.Ile722Thr	rs1800417	0.05	1/8	Missense	0.28 ^d	Likely Benign	PM2, BP4 BP6	3
<i>OCA2</i>	C.1183-4A>G*	-	rs10852218	0.42	6/8	Splice variant	0.0 ^c	Benign	BA1	-
<i>OCA2</i>	C.913C>T*	p.Arg305Trp	rs1800401	0.13	7/8	Missense	0.32 ^d	Benign	BA1	-

^a = Population allele frequency as reported by 1000 Genome Project.AFR
^b = In this study
^c = SpliceAI predicted effect on splicing;
^d = REVEL score predicting the pathogenicity of missense variants
* = Variants excluded from further analysis
¹ = [Kerr et al. \(2000\)](#); [Rimoldi et al. \(2014\)](#);ClinVar-VCV000211768.17
² = [Kamaraj and Purohit \(2014\)](#); ClinVar-VCV000255713.13
³ = [Kamaraj and Purohit \(2014\)](#); [Kerr et al. \(2000\)](#); [Martin et al. \(2017\)](#);ClinVar-VCV000255728.14

The variant NM_000275.3:c.1007C>T was identified in P1 who is heterozygous for the 2.7kb deletion, whilst NM_000275.3:c.2165T>C was identified in P7 who tested negative for the 2.7kb deletion. In silico tools predicted the effect of the variants to be tolerated with overall low REVEL scores (Table 2), as such benign code BP4 was applied.

Upon the assessment of the variant positions (NM_000275.3:c.1007C>T and NM_000275.3:c.2165T>C), neither variant was associated with other pathogenic novel missense variants or same amino acid changes at that position. As such neither PM5 nor PS1 were applied to these variants. These variants were both reported as benign by reputable sources (ClinVar) and in the literature, therefore BP6 was applied. Ultimately, these two variants were classified as likely benign due to ACMG code combination (PM2, BP4, BP6) favouring a more benign contribution.

The variant NM_000275.3:c.1503+5G>A, seen in P3, P5 and P8, has been shown to have functional consequences (Rimoldi et al., 2014), and is reported as clinically significant in ClinVar (ClinVar ID: 211768), PS3 and PP5 were thus respectively applied. The silico tool, SpliceAI, predicts this variant to have a high splice-altering effect with a high donor loss score of 0.76, for this reason, PP3 was applied. Visualisation on IGV confirmed the variant to be true and present in a heterozygous state (Figure 4). PVS1 was not applied as this splice variant is not within ± 2 base pairs of the intron thus not compatible with splice site recommendations. This variant

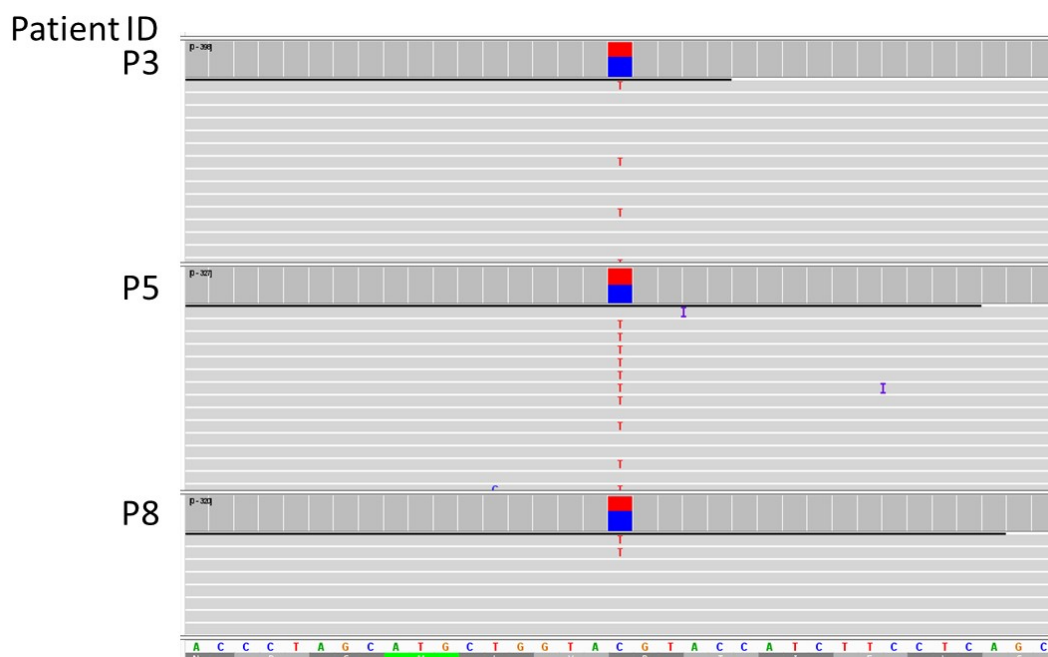


Figure 4: IGV visualisation of NM_000275.3:c.1503+5G>A (rs368124046) variant in P3, P5 and P8. The variant is located in position chr15:28228486 and appears in a heterozygous state in all three individuals.

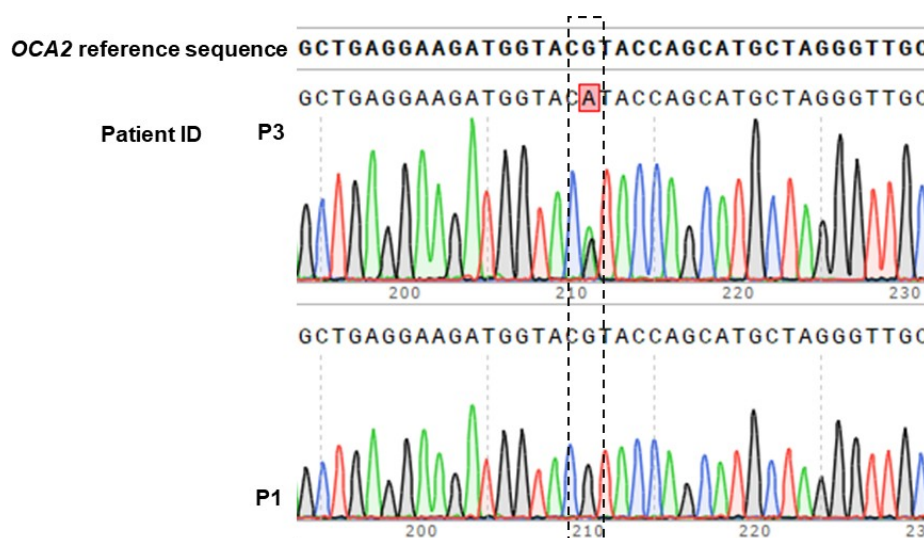


Figure 5: Sanger sequencing validation for NM.000275.3:c.1503+5G>A (rs368124046). Sanger sequencing was performed on the 3500xL genetic analyser and analysed using SnapGene software. Pane 1 labelled as OCA2 reference sequence shows the alignment sequence utilised: GRCh37:ENSG00000104044 assembly from Ensembl. Pane 2 and 3 show the region of interest for Patient ID P3 and P1 respectively. For P3, heterozygosity for the NM.000275.3:c.1503+5G>A variant was seen as an AG double peak at location 15:28228486, with the mutant allele (A) represented by the green peak. Homozygosity for the wildtype allele (GG) was seen for P1, represented by a single black peak. Results for WES were therefore validated.

was classified in this study as likely pathogenic owing to multiple combinations of pathogenicity codes (PS3, PM2, PP3, PP5). This variant was validated using Sanger sequencing (Figure 5). Identification of this variant contributes 37.5% (3/8 unknown chromosomes) to the overall diagnostic yield within this cohort.

3.3 Other pigment genes

HERC2, known for housing an enhancer for the *OCA2* gene, was investigated in individuals (P1, P2, P4, P6, P7) who remained with an unknown second variant causing OCA2. The variant NM_004667.6:c.13272+874T>C (rs12913832), known to reduce *OCA2* expression (Eisenberg, 2020; Visser et al., 2012), was not identified. No variants of interest were detected in the *TYR* gene.

A total of four variants were prioritised in *TYRP1* (Table 3). Two of the variants were found to be known pathogenic variants, NM_000550.3:c.497C>G (p.Ser166Ter) (rs10489413) and NM_000550.3:c.1103del (p.Lys368fs) (rs387906560), and were in a compound heterozygous state in P6. These two variants are well described and are known to be causative of the ROCA phenotype. Substantial evidence for the pathogenicity of NM_000550.3:c.497C>G and NM_000550.3:c.1103del could be gathered through the literature and ClinVar submissions.

Table 3: *TYRP1* (NM_000550.3) variants prioritised using ACMG/AMP guidelines

Gene	Variation		SNP	Allele frequency ^a	Occurrence	Consequence	In silico prediction ^b	Classification	Evidence	Ref
	HGVSc	HGVSp								
<i>TYRP1</i>	C.1557T>G	p.Tyr519Ter	rs41302073	0.05	1/8	Stop gained	-	Benign/Likely benign ^c	PM2, PM6	1
<i>TYRP1</i>	C.977G>A	p.Arg326His	rs16929374	0.02	1/8	Missense	0.07	Likely Benign	PM2, BP1 BP5, BP6	2
<i>TYRP1</i>	C.497C>G	p.Ser166Ter	rs104894130	0	1/8	Stop gained	-	Pathogenic	PVS1, PM2 PP5	3
<i>TYRP1</i>	C.1103del	p.Lys368fs	rs387906560	0	1/8	Frameshift	-	Pathogenic	PVS1, PM2 PP5	4

^a = Population allele frequency as reported by 1000 Genome Project.AFR
^b = REVEL score predicting the pathogenicity of missense variants. In silico predictions for frameshift and stop gain variants were not identified
^c = Classification as noted on ClinVar
¹ = ClinVar - VCV000198743.18
² = ClinVar - VCV000437183.19
³ = (Chiang et al., 2008; Manga et al., 1997); ClinVar-VCV000017594.4
⁴ = (?Manga et al., 1997); ClinVar-VCV000017593.22

The remaining variants, NM_000550.3:c.1557T>G (p.Tyr519Ter) (rs41302073) and NM_000550.3:c.977G>A (p.Arg326His) (rs16929374) were found in P2 and P3 respectively in a heterozygous state. Both NM_000550.3:c.1557T>G and NM_000550.3:c.977G>A were reported as benign/likely benign on ClinVar. Further classification of these variants was not pursued.

3.4 Overall result summation

A total of three clinically significant variants were identified in four individuals, resolving 50% (4/8) cases in this study (Table 4). Variants NM_000550.3:c.1103del and NM_000550.3:c.497C>G identified in the *TYRP1* gene resulted in a diagnosis for P6. Results for this patient are thus considered separately from the rest of the cohort, as this individual does not actually have OCA2. Variant NM_000275.3:c.1503+5G>A was found in 3/7 possible OCA2 patients resulting in a 42.9% OCA2 diagnostic resolution rate. Causative variants and a confirmed diagnosis remain elusive for the remaining four patients (Table 4).

4 DISCUSSION

This study presents the results of exome sequencing analysis of a cohort of eight unrelated individuals of African descent with a reported OCA2 phenotype. A total of three clinically significant variants were identified, with one in *OCA2* and the remaining two in the *TYRP1* gene. Of the eight individuals tested, the second recessive variant was identified in three of them, and a fourth individual was shown to have as opposed to OCA2.

4.1 Variants in *OCA2*

4.1.1 Detection of SNVs

A likely pathogenic splice variant, NM_000275.3:c.1503+5G>A, was identified in three individuals within this cohort. This variant was previously identified in several African individuals and classified as non-pathogenic at the time (Kerr et al., 2000), however, it was later investigated at the functional level and shown to be causative of OCA2 through in vitro analysis (Rimoldi et al., 2014). Variants at this intronic position result in the activation of a cryptic 5' splice site (Buratti et al., 2007; Rimoldi et al., 2014) consequently resulting in aberrant RNA in a gene where loss of function is a known mechanism of disease (Rimoldi et al., 2014). The exact function of the OCA2 protein is not well understood, however, variants that disrupt the function or expression of this protein are known to impair melanin biogenesis thus leading to hypopigmentation (Chen et al., 2002; Hirobe, 2011; Toyofuku et al., 2002).

Table 4: Summary of results and variant profile per case

Code	Gene	Consequence	Genotype	Variant	Zygosity	Phenotype	Confirmed diagnosis
P1	OCA2	deletion -	M/U	2.7kb del unknown	het -	-	-
P2	OCA2	deletion -	M/U	2.7kb del unknown	het -	-	-
P3	OCA2	deletion -	M/M	2.7kb del c.1503+5G>A	het het	presents with albinism	OCA2
P4	OCA2	deletion deletion	M/U	2.7kb del unknown	het -	partial albinism; no nystagmus; medium fair complexion; blackish-brown hair	-
P5	OCA2	deletion -	M/M	2.7kb del c.1503+5G>A	het het	presents with albinism	OCA2
P6	OCA2	deletion	M/U	2.7kb del	het	?albinism	ROCA
	TYRP1	stop_gained frameshift_variant	M/U	c.497C>G (p.Ser166Ter) c.1103del (p.Lys368fs)	het het		
P7	OCA2	-	U/U	unknown	-	clinically pale hair and eye colour; NICU baby	-
P8	OCA2	deletion splice donor	M/M	2.7kb del c.1503+5G>A	het het	-	OCA2

U = unknown variant
M = variant identified
OCA2 transcript = NM_000275.3
TYRP1 transcript = NM_000550.3
het = heterozygous

This variant accounts for 21.4% (3/14) of OCA2 chromosomes within this study, slightly higher than the 18% (14/78 chromosomes) observed in an older study, albeit the sample size was larger than this current study (Kerr et al., 2000). With the exclusion of P6 (confirmed), the diagnostic rate for OCA2 thus increased from 37.5% to 42.9% (3/7) in this cohort.

The absence of functional data for a majority of variants often impairs definitive classification for those variants. The functional characterisation of the NM_000275.3:c.1503+5G>A variant proved to be the pivotal step required to provide a more definitive assessment of the variant's impact (Rimoldi et al., 2014), ultimately leading to its reclassification from non-pathogenic to likely pathogenic in this paper. This thus highlights the importance of studying the functional effect of variants on a gene to better understand their role in disease onset. (Brnich et al., 2020; Richards et al., 2015)

4.1.2 Detection of CNVs

No CNVs, including the 2.7kb deletion, were identified using WES. Although a 50% reduction in coverage for exon 7 was anticipated in P3, P5, and P8, who were heterozygous for the 2.7kb deletion, this was not observed upon visual inspection using IGV. Despite the noticeable reduction in coverage when zoomed out, a mathematical analysis comparing the coverage of P3 and P7 (who tested negative for the 2.7 kb deletion) revealed only a 25.1% difference in coverage (Figure S1), this was not significant enough to confirm the deletion. Other studies investigating the *OCA2* gene using WES and CNV analysis tools different from Ion Reporter have succeeded in identifying large deletions within the gene, however, none have been sensitive enough to detect smaller CNVs (commonly called small insertions or deletions (INDELs)) such as the 2.7kb deletion (Jiang et al., 2024; Shahzad et al., 2017). INDELs are noted to be less than 10kb in size (Mullaney et al., 2010), the Ion Reporter Software reportedly identifies CNVs from 10mb in size (MAN0028321 A.0, Thermo Fisher Scientific). This is important to note as it suggests that the possible presence of smaller CNVs and INDELs is being missed, further highlighting the limitation of CNV analysis using WES.

4.2 Variants in *TYRP1*

Two common *TYRP1* variants (NM_000550.3:c.1103del and NM_000550.3:c.497C>G) were identified in individual P6. The variants are known to be causative of the (classified as OCA3) phenotype (Kromberg et al., 2012; Manga et al., 1997). The additional, heterozygous presence of the 2.7kb deletion in *OCA2* results in a more hypopigmented phenotype with lighter skin and yellow-like hair closely resembling OCA2. This combination phenotype has been described

previously ([Manga et al., 1997](#)). Identification of a case within this cohort highlights the importance of phenotypic clinical details for directive molecular testing.

4.3 The remaining unresolved cases

Although OCA is a visually obvious disorder, phenotypic differences and molecular confirmation of causative variants allow for the distinction between OCA types. Albeit having a family history of albinism, the indistinctive phenotypes described for P4 and P7 highlight the importance of a molecularly confirmed diagnosis.

P4 was heterozygous for the 2.7kb deletion with no other variant identified, similar to two other patients in this study (Table 4), however the inability to identify a second causative variant in the *OCA2* gene does not refute an *OCA2* diagnosis. Partial pigmentation described in P4 (Table 1) is suggestive of a possible milder variant downregulating the expression of the *OCA2* gene, thus altering the phenotype. Variants known to influence *OCA2* gene expression, such as rs12913832 in *HERC2*, could not be investigated in this exome analysis due to their deep intronic location ([Eisenberg, 2020](#); [Sturm et al., 2008](#); [Visser et al., 2012](#)).

For P7, PCR testing confirmed a negative result for the common 2.7kb deletion, and exome analysis did not reveal other potentially causative variants in either *OCA2*, *TYR*, or *TYRP1*. As this individual was a newborn and in the neonatal intensive care unit (NICU), a clinical picture, other than classic OCA2, is suggested. Cases of hypopigmentation that do not arise from variants in associated pigment genes could be linked to syndromic albinism disorders such as Hermansky-Pudlak Syndrome (HPS) ([Merideth et al., 2009](#); [Thomas et al., 2023](#); [Wei and Li, 2013](#)). Diagnosis of syndromic albinism required clinical information that was not available, and beyond the scope of the current study.

4.4 Limitations and future directions

Ultimately, causative variants in four individuals in this cohort remain unknown. This is ascertained by limitations within the sequencing technology used here. The lack of variants identified in these individuals suggests that the variants are in regions undetected by WES. Detection of CNVs using WES data is based on the comparison between sequence read depth ([Mason-Suares et al., 2016](#); [Moreno-Cabrera et al., 2020](#); [Zhao et al., 2013](#)) and an appropriate baseline or reference sequence similar to the study population (African descent). CNV analysis was largely limited in this study by the capability of the analysis software used. Inability to identify the common 2.7kb deletion in known carriers using Ion Reporter software and IGV showed that the analysis of small CNVs/INDELs from WES data was unreliable in this study. Furthermore, alignment to

GRCh37 utilising Ion torrent sequencing posed a limitation on GATK usage due to the immediate unavailability of resources for the aligned build and sequencer used.

Further analysis to be considered for these samples includes whole genome sequencing (WGS) or high-resolution comparative genome hybridization microarray (array-CGH). Several complex rearrangements in *OCA2* have been identified and mapped through the use of array-CGH (Morice-Picard et al., 2014). Variants identified in African individuals included the 2.7kb deletion, exon 5 deletion (2.4kb), and a large deletion including the whole *OCA2* gene (Morice-Picard et al., 2014). On the other hand, WGS would be useful in the investigation of deep intronic regions possibly housing regulatory elements, and also large CNVs missed by detection and analysis techniques used in the current study.

5 CONCLUSION

This study was successful in its aim to investigate variants other than the 2.7kb deletion that are causative of the *OCA2* phenotype in the Southern African population. The splice site variant, NM_000275.3:c.1503+5G>A, identified herewith exhibits a high detective rate and is suggestive of a second common variant in this population. Identification of this mutations will assist in the development of more accurate molecular testing approaches that can improve the correct classification of *OCA2* cases in Africa, as well as provide carrier and prenatal testing for families with a history of *OCA2*.

Future research to determine the frequency of this variant within a larger sample size is required. Findings in this study thus indicate the necessity for further investigation of other variants causative of *OCA2*. The identification of ROCA through molecular diagnosis in this study highlights the importance of precisely describing the phenotype and clinical features experienced by patients for better differential diagnosis.

DATA AVAILABILITY STATEMENT

The datasets utilized 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 (M230778).

AUTHOR CONTRIBUTIONS

TGN designed the study. PMM performed PCR, WES, and Sanger sequencing under the supervision of TGN. PTM and PMM performed bioinformatics analysis of the data. PM prepared the draft manuscript. RAK and TGN provided theoretical guidance. TGN, RAK, and PTM revised the manuscript and read the final manuscript. All the authors have thoroughly reviewed the manuscript and have given their approval for the submitted version.

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 affirm that the study was conducted without any conflicts of interest.

REFERENCES

- Abou Tayoun, A. N., Pesaran, T., DiStefano, M. T., Oza, A., Rehm, H. L., Biesecker, L. G., et al. (2018). Recommendations for interpreting the loss of function PVS1 ACMG/AMP variant criterion. *Hum Mutat* 39, 1517–1524. doi: [10.1002/humu.23626](https://doi.org/10.1002/humu.23626)
- Brnich, S. E., Abou Tayoun, A. N., Couch, F. J., Cutting, G. R., Greenblatt, M. S., Heinen, C. D., et al. (2020). Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med* 12, 3. doi: [10.1186/s13073-019-0690-2](https://doi.org/10.1186/s13073-019-0690-2)
- Buratti, E., Chivers, M., Královičová, J., Romano, M., Baralle, M., Krainer, A. R., et al. (2007). Aberrant 5 splice sites in human disease genes: mutation pattern, nucleotide structure and comparison of computational tools that predict their utilization. *Nucleic Acids Res* 35, 4250–4263. doi: [10.1093/nar/gkm402](https://doi.org/10.1093/nar/gkm402)

- Chen, K., Manga, P., and Orlow, S. J. (2002). Pink-eyed Dilution Protein Controls the Processing of Tyrosinase. *Molecular Biology of the Cell* 13, 1953–1964. doi: [10.1091/mbc.02-02-0022](https://doi.org/10.1091/mbc.02-02-0022)
- Chiang, P.-W., Fulton, A. B., Spector, E., and Hisama, F. M. (2008). Synergistic interaction of the *OCA2* and *OCA3* genes in a family. *Am. J. Med. Genet.* 146A, 2427–2430. doi: [10.1002/ajmg.a.32453](https://doi.org/10.1002/ajmg.a.32453)
- 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. doi: [10.1038/ng0694-176](https://doi.org/10.1038/ng0694-176)
- Eisenberg, M. T. (2020). *Investigation of the OCA2 gene control regions and their possible role in normal pigment variation*. MSc dissertation, University of the Witwatersrand
- Grønskov, K., Ek, J., and Brøndum-Nielsen, K. (2007). Oculocutaneous albinism. *Orphanet J Rare Dis* 2, 43. doi: [10.1186/1750-1172-2-43](https://doi.org/10.1186/1750-1172-2-43)
- Hirobe, T. (2011). How are proliferation and differentiation of melanocytes regulated? *Pigment Cell Melanoma Res* 24, 462–478. doi: [10.1111/j.1755-148X.2011.00845.x](https://doi.org/10.1111/j.1755-148X.2011.00845.x)
- Jiang, B., Zhang, H., Kan, Y., Gao, X., Du, Z., and Liu, Q. (2024). Novel compound heterozygous mutations in *OCA2* gene were identified in a Chinese family with oculocutaneous albinism. *Mol Genet Genomic Med* 12, e2297. doi: [10.1002/mgg3.2297](https://doi.org/10.1002/mgg3.2297)
- Kamaraj, B. and Purohit, R. (2014). Computational Screening of Disease-Associated Mutations in *OCA2* Gene. *Cell Biochem Biophys* 68, 97–109. doi: [10.1007/s12013-013-9697-2](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. doi: [10.1002/\(SICI\)1098-1004\(200002\)15:2<166::AID-HUMU5;3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1098-1004(200002)15:2<166::AID-HUMU5;3.0.CO;2-Z)
- King, R. A., Willaert, R. K., Schmidt, R. M., Pietsch, J., Savage, S., Brott, M. J., et al. (2003). MC1R Mutations Modify the Classic Phenotype of Oculocutaneous Albinism Type 2 (*OCA2*). *Am J Hum Genet.* 73, 638–645. doi: [10.1086/377569](https://doi.org/10.1086/377569)
- Kromberg, J. G. and 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., and Jenkins, T. (2012). Types of albinism in the black Southern Africa population. *East Afr Med J* 89, 20–27. doi: [10.4314/eamj.v89i1](https://doi.org/10.4314/eamj.v89i1)
- Kromberg, J. G. R., Castle, D., Zwane, E. M., and Jenkins, T. (2008). Albinism and skin cancer in Southern Africa. *Clin Genet* 36, 43–52. doi: [10.1111/j.1399-0004.1989.tb03365.x](https://doi.org/10.1111/j.1399-0004.1989.tb03365.x)

- Manga, P., Kromberg, J., Box, N., Sturm, R., Jenkins, T., and 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. doi: [10.1086/301603](https://doi.org/10.1086/301603)
- Manga, P., Kromberg, J. G., Turner, A., Jenkins, T., and Ramsay, M. (2001). In Southern Africa, Brown Oculocutaneous Albinism (BOCA) Maps to the OCA2 Locus on Chromosome 15q: P-Gene Mutations Identified. *Am J Hum Genet.* 68, 782–787. doi: [10.1086/318800](https://doi.org/10.1086/318800)
- Martin, A. R., Lin, M., Granka, J. M., Myrick, J. W., Liu, X., Sockell, A., et al. (2017). An Unexpectedly Complex Architecture for Skin Pigmentation in Africans. *Cell* 171, 1340–1353.e14. doi: [10.1016/j.cell.2017.11.015](https://doi.org/10.1016/j.cell.2017.11.015)
- Mason-Suares, H., Landry, L., and S. Lebo, M. (2016). Detecting Copy Number Variation via Next Generation Technology. *Curr Genet Med Rep* 4, 74–85. doi: [10.1007/s40142-016-0091-4](https://doi.org/10.1007/s40142-016-0091-4)
- Merideth, M. A., Vincent, L. M., Sparks, S. E., Hess, R. A., Manoli, I., O'Brien, K. J., et al. (2009). Hermansky-Pudlak Syndrome in Two African-American Brothers. *Am J Med Genet A* 149A, 987–992. doi: [10.1002/ajmg.a.32757](https://doi.org/10.1002/ajmg.a.32757)
- Montoliu, L., Grønskov, K., Wei, A.-H., Martínez-García, M., Fernández, A., Arveiler, B., et al. (2014). Increasing the complexity: new genes and new types of albinism. *Pigment Cell Melanoma Res* 27, 11–18. doi: [10.1111/pcmr.12167](https://doi.org/10.1111/pcmr.12167)
- Moreno-Cabrera, J. M., del Valle, J., Castellanos, E., Feliubadaló, L., Pineda, M., Brunet, J., et al. (2020). Evaluation of CNV detection tools for NGS panel data in genetic diagnostics. *Eur J Hum Genet* 28, 1645–1655. doi: [10.1038/s41431-020-0675-z](https://doi.org/10.1038/s41431-020-0675-z)
- 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. doi: [10.1111/pcmr.12173](https://doi.org/10.1111/pcmr.12173)
- Mullaney, J. M., Mills, R. E., Pittard, W. S., and Devine, S. E. (2010). Small insertions and deletions (INDELs) in human genomes. *Hum Mol Genet* 19, R131–R136. doi: [10.1093/hmg/ddq400](https://doi.org/10.1093/hmg/ddq400)
- Oetting, W. S., Garrett, S. S., Brott, M., and King, R. A. (2005). P gene mutations associated with oculocutaneous albinism type II (OCA2). *Hum. Mutat.* 25, 323–323. doi: [10.1002/humu.9318](https://doi.org/10.1002/humu.9318)
- Puri, N., Durham-Pierre, D., Aquaron, R., Lund, P. M., King, R. A., and Brilliant, M. H. (1997). Type 2 oculocutaneous albinism (OCA2) in Zimbabwe and Cameroon: distribution of the 2.7-kb deletion allele of the P gene. *Hum Genet* 100, 651–656. doi: [10.1007/s004390050568](https://doi.org/10.1007/s004390050568)
- 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. doi: [10.1038/gim.2015.30](https://doi.org/10.1038/gim.2015.30)
- Rimoldi, V., Straniero, L., Asselta, R., Mauri, L., Manfredini, E., Penco, S., et al. (2014). Functional characterization of two novel splicing mutations in the OCA2 gene associated with oculocutaneous albinism type II. *Gene* 537, 79–84. doi: [10.1016/j.gene.2013.11.102](https://doi.org/10.1016/j.gene.2013.11.102)
- Robinson, J. T., Thorvaldsdóttir, H., Wenger, A. M., Zehir, A., and Mesirov, J. P. (2017). Variant Review with the Integrative Genomics Viewer. *Cancer Res* 77, e31–e34. doi: [10.1158/0008-5472.CAN-17-0337](https://doi.org/10.1158/0008-5472.CAN-17-0337)
- Shahzad, M., Yousaf, S., Waryah, Y. M., Gul, H., Kausar, T., Tariq, N., et al. (2017). Molecular outcomes, clinical consequences, and genetic diagnosis of Oculocutaneous Albinism in Pakistani population. *Sci Rep* 7, 44185. doi: [10.1038/srep44185](https://doi.org/10.1038/srep44185)
- Spritz, R. A., Fukai, K., Holmes, S. A., and Luande, J. (1995). Frequent Intragenic Deletion of the P Gene in Tanzanian Patients with Type 11 Oculocutaneous Albinism (OCA2). *Am. J. Hum. Genet.*
- Stevens, G., Ramsay, M., and 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. doi: [10.1007/s004390050400](https://doi.org/10.1007/s004390050400)
- Stevens, G., van Beukering, J., Jenkins, T., and 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.*
- Sturm, R. A., Duffy, D. L., Zhao, Z. Z., Leite, F. P., Stark, M. S., Hayward, N., et al. (2008). A Single SNP in an Evolutionary Conserved Region within Intron 86 of the HERC2 Gene Determines Human Blue-Brown Eye Color. *Am J Hum Genet.* 82, 424–431. doi: [10.1016/j.ajhg.2007.11.005](https://doi.org/10.1016/j.ajhg.2007.11.005)
- Thomas, M. G., Zippin, J., and Brooks, B. P. (2023). Oculocutaneous Albinism and Ocular Albinism Overview. In *GeneReviews®*, eds. M. P. Adam, J. Feldman, G. M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, and A. Amemiya (Seattle (WA): University of Washington, Seattle)
- Toyofuku, K., Valencia, J. C., Kushimoto, T., Costin, G.-E., Virador, V. M., Vieira, W. D., et al. (2002). The Etiology of Oculocutaneous Albinism (OCA) Type II: The Pink Protein Modulates the Processing and Transport of Tyrosinase. *Pigment Cell Res* 15, 217–224. doi: [10.1034/j.1600-0749.2002.02007.x](https://doi.org/10.1034/j.1600-0749.2002.02007.x)

- Visser, M., Kayser, M., and Palstra, R.-J. (2012). *HERC2* rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the *OCA2* promoter. *Genome Res.* 22, 446–455. doi: [10.1101/gr.128652.111](https://doi.org/10.1101/gr.128652.111)
- Wei, A.-H. and Li, W. (2013). Hermansky-Pudlak syndrome: pigmentary and non-pigmentary defects and their pathogenesis. *Pigment Cell Melanoma Res* 26, 176–192. doi: [10.1111/pcmr.12051](https://doi.org/10.1111/pcmr.12051)
- Zhao, M., Wang, Q., Wang, Q., Jia, P., and Zhao, Z. (2013). Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinformatics* 14, S1. doi: [10.1186/1471-2105-14-S11-S1](https://doi.org/10.1186/1471-2105-14-S11-S1)

6 APPENDICES

Appendix A: Supplementary Material

Genotyping the 2.7kb deletion

The PCR reaction included 1 μ l of each primer (MHB71, MHB72, and MHB107), 1 μ l of 100 ng genomic DNA, 12.5 μ l DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific, South Africa), and 8.5 μ l of nuclease-free water to a total volume of 25 μ l. PCR amplification was carried out on a Bio-Rad T100 Thermal Cycler using the following conditions: denaturation at 95°C for one minute, annealing at 60°C for one minute, synthesis at 72°C for one minute, repeated 35 times. PCR products were separated on a 3% agarose gel and ethidium bromide staining was used for visualisation under ultraviolet light using the Omega Fluor™ Gel documentation system (Vacutec, South Africa).

Variant validation using Sanger sequencing

The PCR reaction included 0.8 μ l of each primer (Table S1) at 10pmol, 1 μ l of 100 ng genomic DNA, 10 μ l DreamTaq Green PCR Master Mix (2X) (Thermo Fisher Scientific, South Africa), and made up to a total volume of 20 μ l using nuclease-free water. PCR amplification was carried out on a Bio-Rad T100 Thermal Cycler using the following conditions: denaturation at 95°C for 30 seconds, annealing at 54°C for 30 seconds, synthesis at 72°C for 30 seconds, repeated 35 times. The sample was maintained at a temperature of 72°C for 5 minutes, to ensure full extension. 5 μ l of PCR products were separated on a 3% agarose gel and ethidium bromide staining was used for visualisation under ultraviolet light using the Omega Fluor™ Gel documentation system (Vacutec, South Africa).

The PCR products were cleaned using 6 μ l ExoSAP-IT™ Express (Thermo Fisher Scientific, South Africa), followed by incubation at 37°C and 80°C for 15 min each. Cycle sequencing reaction was performed using 5 μ l of the purified PCR product in a reaction mix containing 2 μ l Bigdye terminator v3.1 Ready Reaction Mix (Applied Biosystems, USA), 1 μ l 5x Bigdye sequencing buffer (Applied Biosystems, USA), 1 μ l of 10pmol forward primer (Inqaba Biotechnical Industries, SA) and 1 μ l of nuclease-free water.

Table S1: PCR primer sequences designed to amplify the NM.000275.3:c.1503+5G>A variant region

Primer	Sequence (5'→3')	Length(bp)
rs368124046_F	AAC AGT CAC CTA ACA TCC CA	20
rs368124046_R	CCT CCC TTA TAC GAG CAA GC	20

Post-cycle sequencing clean-up was performed using the BigDye XTerminator™ Purification Kit as per manufacturer's instruction. 45 µl SAM™ Solution and 10ul BigDye XTerminator™ Solution were added to 10µl of the cycle sequencing product. The plate was subsequently sealed using MicroAmp™ Clear Adhesive Film (Applied Biosystems, USA), vortexed for 30 minutes on a 96-well microplate mixer, and finally centrifuged for 2 minutes. Sequencing was conducted on the ABI 3500xl genetic analyzer and results were analysed using Snapgene software.

Detection of CNVs

Percentage of reduced coverage was calculated using the following formula:

$$\% \text{Coverage reduced} = 100 - \left[\frac{\text{Total count of heterozygous sample}}{\text{Total count of negative sample}} * 100 \right]$$

Coverage for P3 was thus found to be reduced by 25.1%, but not by approximately 50% as expected for a known 2.7kb deletion carrier (Figure S1).

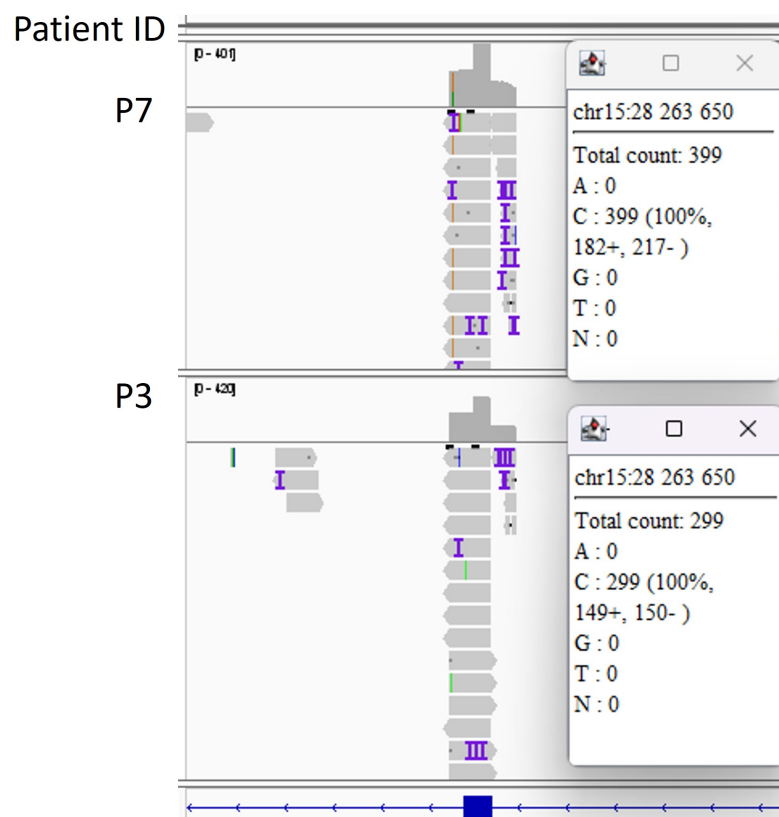


Figure S1: IGV visualization of the 2.7kb deletion region. The 2.7kb deletion is mapped at chr15:28262863-28265786 and encompasses a deletion of exon 7. Pane 1 and 2 visualise the coverage (total count) of P7 (negative for the 2.7kb deletion) and P3 (heterozygous for the 2.7kb deletion) respectively for the same position. A 25.1% reduction in coverage is seen for P3 compared to P7.

Appendix B: Human Research Ethics Committee (Medical), University of the Witwatersrand (M230778)



R49 Ms P Mmbi

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M230778

NAME:
(Principal Investigator)

Ms P Mmbi

DEPARTMENT:

School of Pathology
Division of Human Genetics
Medical School
University

PROJECT TITLE:

*Investigating the genetic cause of oculocutaneous albinism
type 2 in individuals of African descent through exome
sequencing*

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>

SUPERVISOR:

Drs T Ngcungcu, P Mpangase and R Kerr

APPROVED BY:

Professor P Ruff, Chairperson, HREC (Medical)

DATE OF APPROVAL:

2023/08/14

EXPIRY DATE:

2028/08/13

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

DECLARATION OF INVESTIGATORS

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

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. **I agree to submit a yearly progress report.** When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in July and therefore reports and re-certification will be due in the month of July each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).


Signature of Principal Investigator

15/08/2023
Date

Appendix C: Human Research Ethics Committee (Medical), University of the Witwatersrand (M210203)



R49 Dr T Ngcungcu

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

NAME: Dr T Ngcungcu
(Principal Investigator)

DEPARTMENT: School of Pathology
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*

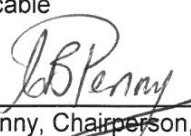
DATE CONSIDERED: 2021/02/26

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)

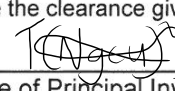
DATE OF APPROVAL: 2021/08/10

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



CANDIDATE'S SURNAME: [Please print]	Mmbi	FIRST NAME/S:	Phophi	STUDENT NUMBER:	1647122
CURRENT QUALIFICATIONS: BSc (Hon) Medical Sciences					
TEL:	CELL: 073 418 3794	E-MAIL:	1647122@students.wits.ac.za	FAX:	
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: MSc (Med) Genomic Medicine					
PART-TIME OR FULL-TIME: Full Time					
FIRST REGISTERED FOR THIS DEGREE:	TERM : 1		YEAR: 2023		
DEPARTMENT: Human Genetics, School of Pathology					
TITLE OF PROPOSED RESEARCH: Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of African descent through exome sequencing					
CANDIDATE'S SIGNATURE:				DATE: 23/04/2023	
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 Phelelani T. Mpangase				% Supervision = 25%	
SUPERVISOR'S QUALIFICATIONS : PhD					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: Phelelani.Mpangase@wits.ac.za					
SUPERVISOR 3 (NAME & SURNAME): Dr Robyn Kerr				% Supervision = 25%	
SUPERVISOR'S QUALIFICATIONS: PhD					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: Robyn.Kerr@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]					

WITS ETHICS NOT REQUIRED: WITS ETHICS PENDING: WITS ETHICS APPROVED: (circle appropriate symbol)*	Yes No Yes No Yes No	IF ☐ 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

Research Proposal

Investigating the genetic cause of oculocutaneous albinism type 2 in individuals of
African descent through exome sequencing

by

Phophi M. Mmbi

1647122

Master of Science in Medicine in Genomic Medicine

Supervisor 1

Dr Thandiswa G. Ngcungcu

PhD (Wits), Lecturer, Division of Human Genetics, University of the Witwatersrand

Supervisor 2

Dr Robyn A. Kerr

PhD (Wits), Lecturer, Division of Human Genetics, University of the Witwatersrand

Supervisor 3

Dr Phelelani T. Mpangase

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

Table of Contents

1. Background..... 1

2. Aim..... 3

3. Study Objectives..... 4

4. Methods..... 4

5. Data Analysis..... 6

7. Ethics..... 7

8. Timing..... 7

9. Funding..... 8

10. Possible limitations..... 8

11. References..... 9

1. Background

Albinism is a group of inherited, autosomal recessive, hypopigmentation disorders. There are two types of albinism: (i) ocular albinism, a rarer form mainly affecting the eyes; and (ii) oculocutaneous albinism (OCA), a common trifactor form affecting the eyes, hair and skin (reviewed in Montoliu et al., 2014). OCA shows locus heterogeneity and presents as seven types, differentiated by phenotype and genetic cause. OCA2 is caused by mutations in the OCA2 gene, which encodes a transmembrane protein involved in the processing and regulation of melanosomal proteins (Chen et al., 2002; Toyofuku et al., 2002). The clinical features of OCA2 include hypopigmentation of the skin, hair, and eyes, as well as visual impairment and predisposition to skin cancer (Kromberg et al., 2008). Individuals with OCA2 have light- coloured skin and hair, and their eyes may appear blue-grey, or brown. The incidence of OCA2 in black South Africans is 1/3900 births, but the prevalence varies across sub-Saharan Africa (Stevens et al., 1997).

The most common cause of OCA2 in sub-Saharan Africans is a 2.7kb deletion mutation, which includes the deletion of exon 7 of the OCA2 gene (Durham-Pierre et al., 1994; Stevens et al., 1995). This mutation accounts for 78% of OCA2 chromosomes, leaving 22% of mutations unidentified in this population group (Stevens et al., 1997). Therefore, many individuals with OCA2 are homozygous for the 2.7kb deletion, but about 38% of individuals with OCA2 are either heterozygous (30%) or have completely unknown (8%) genotypes (Stevens et al., 1997). Thus, a large proportion of disease-causing OCA2 mutations remain to be identified (Stevens et al., 1997).

A few studies have been conducted across Africa in attempts to identify other mutations that cause or contribute to the OCA2 phenotype. In Tanzania, Spritz et al., (1995) identified two deleterious novel mutations in a compound heterozygous state with the common 2.7kb deletion: (i) a splice site variant (IVS21-2A>G) and (ii) missense variant (K614N). Their study further identified a Nigerian couple who were carriers for novel frameshift (V726- L727delGTCCT) and missense (T404M) variants in the OCA2 gene that were assumed to have caused the OCA2 phenotype in their affected child (Spritz et al., 1995). In the Sub-Saharan region, Kerr et al., (2000) identified one missense (A334V) and three frameshift (614delA, 683insT, and 727insG) mutations that had a likely causative contribution to an OCA2 phenotype. These studies suggest that a second common mutation does not exist and that 'other' OCA2 mutations are private. To our knowledge, no other recent studies have been conducted to try and identify other OCA2 causing mutations in this population.

In more recent years, several studies have investigated the link between polymorphisms in modifier genes and their role in the differing OCA2 phenotypic expression in certain populations (Chiang et al., 2008, Johanson et al., 2010; King et al., 2003). These studies looked at genes such as HERC2 and MC1R, and drew conclusions on the effect of variants in these genes on hair, skin and eye colour for their populations. In the study conducted on Polynesian individuals it was revealed that MC1R polymorphisms did not affect or contribute to the OCA2 phenotype, confirming that effects of this gene are population specific (Johanson et al., 2010). In the study conducted by Chiang and colleagues (2008), they investigated modified hair colour in an OCA2 patient of Hispanic origin and found no involvement of polymorphisms in the MC1R gene. Instead, they found a mutation in the TYRP1 gene (known to cause another type of OCA, namely OCA3, also autosomal

recessive) and this suggested that variants in OCA genes can have a synergistic interaction with each other to influence a phenotype (Chiang et al., 2008).

Crawford et al., (2017), investigating normal pigment variation, showed that multiple genes are associated with pigment variation in the African population. These genes include the OCA2, HERC2 and other genes involved in other types of OCA (Crawford et al., 2017). The HERC2 gene lies immediately next to the OCA2 gene on chromosome 15q and homozygosity for a SNP (rs1800404) found in HERC2 was shown to reduce OCA2 function by ~40%. While this is not sufficient to produce an OCA2 phenotype, it may have significant effect if present together with a classic OCA2 mutation. If ~29% of African skin pigmentation can be attributed to the additive effect of four genetic regions (OCA2/HERC2 included) (Crawford et al., 2017), is it not possible that interaction between mutations in these genes could result in a hypopigmented phenotype?

A number of suggestions which could be used to explain the large proportion of OCA2 mutations that remain unidentified have been described: (i) limitations in the ability of older testing techniques previously used (Kerr et al., 2000; Oetting et al., 2005); (ii) missing mutations in intronic regions not covered by the testing approach (Kerr et al., 2000; Oetting et al., 2005); (iii) possible epistasis between known pigment genes (Chiang et al., 2008); (iv) presence of mutations in the promoter or enhancer regions of the gene (Oetting et al., 2003) and (v) unknown functionality of splicing mutations (Kerr et al., 2000; Preising et al., 2011; Rimoldi et al., 2014). Current advancements in technology, sequencing techniques and bioinformatic tools can now be used to interrogate genes more fully, and to interrogate several genes simultaneously.

2. Aim

This study thus aims to investigate pigment genes, primarily the OCA2 gene, using exome sequencing, in order to identify possible 'other' mutations in individuals who tested heterozygous for the common OCA2 2.7kb deletion.

3. Study Objectives

1. To perform whole exome sequencing on eight individuals presenting with OCA2
2. To select genes associated with skin pigmentation (through literature search) for analysis only if no mutations are found in the OCA2 gene, otherwise prioritizing the examination of the OCA2 gene
3. To perform bioinformatic analyses to:
 - Identify single nucleotide variants (SNVs) and small insertions/deletions (INDELs) in the OCA2 gene, or other selected genes associated with skin pigmentation if no mutations are found in the OCA2 gene
 - Annotate the identified variants using Ensembl's Variant Effect Predictor (VEP) and other relevant tools
 - Filter and prioritise variants of interest
4. To perform selective variant validation

4. Methods

4.1 Sample selection

The study sample will comprise of eight de-identified, residual, DNA samples obtained through routine diagnostic testing at the Division of Human Genetics, National Health Laboratory Service, Braamfontein. These samples will be from individuals of African descent with OCA2 who were found to be heterozygous for the 2.7kb deletion in the OCA2 gene.

4.2 Whole exome sequencing (WES)

Genetic data will be produced from whole exome sequencing (WES) using the Ion Torrent™ Ion S5 (in-house within the Division of Human Genetics). The advantages of using WES in this study are: (i) this is a massive parallel sequencing technique that will enable us to simultaneously generate data from the OCA2 gene as well as other pigment genes; (ii) OCA2 is a large gene that consists of 23 coding exons. To perform WES would be less time consuming and more cost effective as compared to performing Sanger Sequencing on each exon individually.

The WES workflow will require sample preparation (DNA purification and quality control), library preparation (exome capturing), and sequencing. Data derived from WES will be analysed as described in Section 5.

4.2.1 Sample quality control

Sample quality will be determined using the Qubit™ dsDNA HS Assay Kit with the Qubit™ 3.0 Fluorometer. The integrity of the DNA will then be assessed through agarose gel electrophoresis.

4.2.2. Library preparation

Library preparation will be conducted using the Ion Ampliseq Exome RDY kit. In this process the DNA samples will be amplified, digested and ligated to IonCode™ Barcode Adapters to form the barcoded libraries. The library will thereafter be amplified and finally purified. The Qubit™ 3.0 Fluorometer will be utilised to quantify the concentration of the library.

4.2.3. Sequencing

Barcoded exome libraries will be pooled and will be loaded onto an Ion 540™ Chip, and sequenced using the Ion Torrent™ Ion S5. Sequencing data will be saved as a variant call file (VCF) and will be used as described in Section 5.

4.2.4 Quality Control

The sequencing run will be assessed by evaluating the report metrics (generated by the Ion Torrent software). A reference coverage of >50x; ISP loading of >80%; enrichment of ~100% and a clonality of >70% will be indicative and accepted as good quality for the exome sequencing. Quality assessment using the Integrative Genomics Viewer (IGV) (Robinson et al., 2017) will be conducted for each prioritized variant to further assess the depth of coverage, determine zygosity and exclude artefacts.

5. Data Analysis

Using the VCF generated by the Ion Torrent™ Ion S5 system, the identified variants will be annotated using the Ensembl Variant Effect Predictor (VEP) (McLaren et al., 2016). Annotated variants will be filtered and prioritised following a tiered approach based on: (i) the gene and its curated transcript; (ii) a minor allele frequency of less than 0.05 using the 1000 Genomes Project; (iii) predicted consequence to resultant protein such as frameshift, missense and splice site variants; and (iv) effect of the variant using in silico prediction tools. Prioritised variants will then be classified based on the American College of Medical Genetics and Genomics (ACMG) guidelines for variant interrogation (Richards et al., 2015) and pathogenic and likely pathogenic variants will be assessed further.

6. Variant validation

Variants classified as pathogenic will be validated through Sanger sequencing. Variant validation will be conducted on the condition that time permits. This will involve amplification of DNA sequences through a polymerase chain reaction, followed by amplicon purification and cycle sequencing, and lastly capillary electrophoresis on the ABI 3500xl Genetic Analyzer. If true, these variants will decrease the proportion of unknown or unidentified variants that cause OCA2 and contribute to future testing strategies.

7. Ethics

This research project is a sub-study under the “The identification of causal mutations of classic oculocutaneous albinism type 2 (OCA2) and brown albinism in individuals of African descent” study (PI - Thandiswa Ngcungcu, ethics number M210203). Permission to use the residual samples has been granted by the Division of Human Genetics Head of department (Prof Amanda Krause). Ethics has been applied for and is pending acceptance of this proposal.

8. Timing

[illegible]

9. Funding

Running costs for this study are estimated to be approximately R51 000. The cost of which are illustrated in Table 1 below. This project will be funded by Dr Thandiswa Ngcungcu who received an MRC SIR Grant (2021) for the parent project titled “The identification of causal mutations of oculocutaneous albinism type 2 (OCA2) in individuals of African descent”. All reagents and consumables required for this have been purchased.

Table 1: Cost of exome sequencing

Description	Price (R)
General consumables	1 000.00
Ion Ampliseq Exome RDY kit	20 415.50
Ion 540 chip kit	15 478.75
Ion 540 KIT-chef 8 RXNS	8 814.50
Validation	5 000.00
Total amount	50 708.75

10. Possible limitations

Limitations in this study are based on technique used, sample size and analysis method. Using WES, we are unable to investigate variants found in deep intronic regions and those that are in regulatory elements outside of the exon. This technique reduces the chances of identifying large INDELs and is prone to errors around repetitive sequences.

The use of a small sample size reduces the chances of identifying a possible common mutation. A larger sample size is required in order to accurately estimate the occurrence of a particular variant in the population.

The identification of a Variant of Uncertain Significance (VUS) in the African population may be difficult to resolve due to the diverse nature of the population group and the limited number of studies originating from this population.

11. References

Chen, K., Manga, P. and Orlow, S.J. (2002) 'Pink-eyed Dilution Protein Controls the Processing of Tyrosinase', *Molecular Biology of the Cell*, 13(6), 953–1964.

Chiang, P.-W., Fulton, A.B., Spector, E., Hisama, F.M. (2008) 'Synergistic interaction of the OCA2 and OCA3 genes in a family', *American Journal of Medical Genetics Part A*, 146A (18), 2427–2430.

Crawford, N.G., Kelly, D.E., Hansen, M.E.B., Beltrame, M.H., Fan, S., Bowman, S.L., et al. (2017) 'Loci associated with skin pigmentation identified in African populations', *Science*, 358(6365), p. eaan8433.

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(2), 176–179.

Harding, R.M., Healy, E., Ray, A.J., Ellis, N.S., Flanagan, N., Todd, C., et al. (2000) 'Evidence for Variable Selective Pressures at MC1R', *The American Journal of Human Genetics*, 66(4), 1351–1361.

Johanson, H.C., Chen, W., Wicking, C., Sturm, R.A. (2010) 'Inheritance of a novel mutated allele of the OCA2 gene associated with high incidence of oculocutaneous albinism in a Polynesian community', *Journal of Human Genetics*, 55(2), 103–111.

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', *Human Mutation*, 15(2), 166–172.

King, R.A., Willaert, R.K., Schmidt, R.M., Pietsch, J., Savage, S., Brott, M.J., et al. (2003) 'MC1R Mutations Modify the Classic Phenotype of Oculocutaneous Albinism Type 2 (OCA2)', *The American Journal of Human Genetics*, 73(3), 638–645.

Kromberg, J.G.R., Castle, D., Zwane, E.M., Jenkins, T. (2008) 'Albinism and skin cancer in Southern Africa', *Clinical Genetics*, 36(1), 43–52.

Lee, S.-T., Nicholls, R.D., Jong, M.T.C., Fukai, K., Spritz, R.A. (1995) 'Organization and sequence of the human P gene and identification of a new family of transport proteins', *Genomics*, 26(2), 354–363.

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(1), p. 122.

Montoliu, L., Grønskov, K., Wei, A.-H., Martínez-García, M., Fernández, A., Arveiler, B., et al. (2014) 'Increasing the complexity: new genes and new types of albinism', *Pigment Cell & Melanoma Research*, 27(1), 11–18.

Oetting, W.S., Fryer, J.P., Shriram, S., King, R.A. (2003) 'Oculocutaneous Albinism Type 1: The Last 100 Years', *Pigment Cell Research*, 16(3), 307–311.

Oetting, W.S., Garrett, S.S., Brott, M., King, R.A. (2005) 'P gene mutations associated with oculocutaneous albinism type II (OCA2)', *Human Mutation*, 25(3), 323–323.

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, 939–948.

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

Rimoldi, V., Straniero, L., Asselta, R., Mauri, L., Manfredini, E., Penco, S., et al. (2014) 'Functional characterization of two novel splicing mutations in the OCA2 gene associated with oculocutaneous albinism type II', *Gene*, 537(1), 79–84.

Robinson, J.T., Thorvaldsdóttir, H., Wenger, A.M., Zehir, A., Mesirov, J.P. (2017) 'Variant Review with the Integrative Genomics Viewer', *Cancer Research*, 77(21), e31–e34

Simeonov, D.R., Wang, X., Wang, C., Sergeev, Y., Dolinska, M., Bower, M., et al., (2013) 'DNA Variations in Oculocutaneous Albinism: An Updated Mutation List and Current Outstanding Issues in Molecular Diagnostics', *Human Mutation*, 34(6), 827–835.

Spritz, R.A., Fukai, K., Holmes, S.A., Luande, J. (1995) 'Frequent Intragenic Deletion of the P Gene in Tanzanian Patients with Type 11 Oculocutaneous Albinism (OCA2)', *Am. J. Hum. Genet.* 56(6), 1320-1323.

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(3), 586-591

Stevens, G., Ramsay, M. and Jenkins, T. (1997) 'Oculocutaneous albinism (OCA2) in sub-Saharan Africa: distribution of the common 2.7-kb P gene deletion mutation', *Human Genetics*, 99(4), 523–527.

Toyofuku, K., Valencia, J.C., Kushimoto, T., Costin, G.-E., Virador, V.M., Vieira, W.D., et al. (2002) 'The Etiology of Oculocutaneous Albinism (OCA) Type II: The Pink Protein Modulates the Processing and Transport of Tyrosinase', *Pigment Cell Research*, 15(3), 217–224.

Appendix E: Turnitin Report and Plagiarism Declaration

Phophi_Final_Research_Report-1.pdf

ORIGINALITY REPORT

13%

SIMILARITY INDEX

10%

INTERNET SOURCES

10%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to University of Witwatersrand

Student Paper

2%

2

www.frontiersin.org

Internet Source

2%

3

ebin.pub

Internet Source

1%

4

hdl.handle.net

Internet Source

<1%

5

L. Campbell, J. Fredericks, K. Mathivha, P. Moshesh et al. "The implementation and utility of clinical exome sequencing in a South African infant cohort", Frontiers in Genetics, 2023

Publication

<1%

6

www.vipbg.vcu.edu

Internet Source

<1%

7

worldwidescience.org

Internet Source

<1%

8	Rumi Higuchi, Taichiro Goto, Yosuke Hirotsu, Yujiro Yokoyama et al. "Primary Driver Mutations in GTF2I Specific to the Development of Thymomas", Cancers, 2020 Publication	<1 %
9	rcastoragev2.blob.core.windows.net Internet Source	<1 %
10	"Cassidy and Allanson's Management of Genetic Syndromes", Wiley, 2021 Publication	<1 %
11	Khordadpoor-Deilamani, Faravareh, Mohammad Taghi Akbari, Morteza Karimipoor, and Gholam Reza Javadi. "Homozygosity mapping in albinism patients using a novel panel of 13 STR markers inside the nonsyndromic OCA genes: introducing 5 novel mutations", Journal of Human Genetics, 2016. Publication	<1 %
12	Submitted to University of Salford Student Paper	<1 %
13	onlinelibrary.wiley.com Internet Source	<1 %
14	Oswald, Cody Joe. "A Multi-Omic Approach to Evaluating Soybean Iron Deficiency Chlorosis in the Field.", The University of Nebraska - Lincoln, 2020	<1 %

15

academic.oup.com

Internet Source

<1 %

16

pubmed.ncbi.nlm.nih.gov

Internet Source

<1 %

17

www.mdpi.com

Internet Source

<1 %

18

www.ncbi.nlm.nih.gov

Internet Source

<1 %

19

"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 %

20

Beilei Jiang, Hua Zhang, Yuling Kan, Xueping Gao, Zhaoli Du, Quan Liu. " Novel compound heterozygous mutations in gene were identified in a Chinese family with oculocutaneous albinism ", Molecular Genetics & Genomic Medicine, 2023

Publication

<1 %

21

Robyn Kerr. "Identification of P gene mutations in individuals with oculocutaneous albinism in sub-Saharan Africa", Human Mutation, 02/2000

<1 %

22

ajod.org
Internet Source

<1 %

23

core.ac.uk
Internet Source

<1 %

24

Hygon Mutavhatsindi, Gian D. van der Spuy, Stephanus T. Malherbe, Jayne S. Sutherland et al. "Validation and Optimization of Host Immunological Bio-Signatures for a Point-of-Care Test for TB Disease", *Frontiers in Immunology*, 2021
Publication

<1 %

25

epdf.pub
Internet Source

<1 %

26

www.science.gov
Internet Source

<1 %

27

Annagiusi Gargiulo, Francesco Testa, Settimio Rossi, Valentina Di Iorio et al. "Molecular and Clinical Characterization of Albinism in a Large Cohort of Italian Patients", *Investigative Ophthalmology & Visual Science*, 2011
Publication

<1 %

28

Brilliant, Murray H.. "Albinism in Africa: a medical and social emergency", *International Health*, 2015.
Publication

<1 %

29	Tedder, Matthew L.. "A Healthcare Genetics Approach to Genetic Testing for Hereditary Breast Cancer", Clemson University, 2023 Publication	<1 %
----	---	------

30	Valeria Rimoldi, Letizia Straniero, Rosanna Asselta, Lucia Mauri et al. "Functional characterization of two novel splicing mutations in the OCA2 gene associated with oculocutaneous albinism type II", Gene, 2014 Publication	<1 %
----	---	------

31	citeseerx.ist.psu.edu Internet Source	<1 %
----	--	------

32	docksci.com Internet Source	<1 %
----	--	------

33	ec.bioscientifica.com Internet Source	<1 %
----	--	------

34	epdf.tips Internet Source	<1 %
----	--	------

35	erc.bioscientifica.com Internet Source	<1 %
----	--	------

36	healthdocbox.com Internet Source	<1 %
----	--	------

37	ouci.dntb.gov.ua Internet Source	<1 %
----	--	------

38	people.bath.ac.uk Internet Source	
----	--	--

<1 %

39

www.biorxiv.org

Internet Source

<1 %

40

Advances in Experimental Medicine and Biology, 2014.

Publication

<1 %

41

Jackson Gao, Leera D'Souza, Keith Wetherby, Christian Antolik, Melissa Reeves, David R. Adams, Santa Tumminia, Xinjing Wang. "Retrospective analysis in oculocutaneous albinism patients for the 2.7kb deletion in the OCA2 gene revealed a co-segregation of the controversial variant, p.R305W", Cell & Bioscience, 2017

Publication

<1 %

42

Richard A. King, C. Gail Summers. "Albinism and Hermansky-Pudlak Syndrome", Wiley, 2005

Publication

<1 %

43

"Photosensitivity and Photoreactions in Pediatric Skin of Color", Pediatric Skin of Color, 2015.

Publication

<1 %

44

Monika B. Dolinska, Nicole J. Kus, S. Katie Farney, Paul T. Wingfield, Brian P. Brooks, Yuri V. Sergeev. "Oculocutaneous albinism type 1:

<1 %

link between mutations, tyrosinase conformational stability, and enzymatic activity", Pigment Cell & Melanoma Research, 2017

Publication

Exclude quotes Off

Exclude bibliography On

Exclude matches Off

Appendix F: Frontiers in Genetics Author guidelines

Author guidelines

General standards

Article type

Frontiers requires 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 seen from the 'For authors' menu on every Frontiers journal page. Please pay close attention to the word count limits.

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.

Manuscript length

Frontiers encourages the authors 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.

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, Frontiers recommends 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 Frontiers is pleased to recommend the language-editing service provided by our external partner Editage to authors who believe their manuscripts would benefit from professional editing. 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 Frontiers recommends 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, Frontiers recommends authors to consult the Chicago Manual of Style.

Search engine optimization (SEO)

There are a few simple ways to maximize your article's discoverability. Follow the steps below to improve the search results of your article:

- 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.

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.

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 or 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.

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'

The running title should be a maximum of five words in length.

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.

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.

Keywords

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

Text

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.

Nomenclature

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.

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.

Acknowledgments

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.

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.

Figure and table guidelines

CC-BY license

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.

Figure requirements and style guidelines

Frontiers requires 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. Kindly 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 1 column (85 mm) or 2 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.

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.

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

Please note that 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.

Accessibility

Frontiers encourages authors to make the figures and visual elements of their articles accessible for the visually impaired. An 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

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, PowerPoint presentations, etc.), 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.

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.

References

Frontiers' journals use one of two reference styles, either Harvard (author-date) or Vancouver (numbered). Please check our help center to find the correct style for the journal to which you are submitting.

- All citations 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
- 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 clearly reference the original source
- Preprints can be cited as long as 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)

Many Frontiers journals use the Harvard referencing system; to find the correct reference style and resources for the journal you are submitting to, please visit our help center. Reference examples are found below, for more examples of citing other documents and general questions regarding the Harvard reference style, please refer to the Chicago Manual of Style.

In-text citations

- For works by a single author, include the surname, followed by the year
- For works by two authors, include both surnames, followed by the year
- For works by more than two authors, include only the surname of the first author followed by et al., followed by the year
- For humanities and social sciences articles, include the page numbers.

Reference 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.

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

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.

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

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)

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

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>

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

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

Vancouver reference style (numbered)

Many Frontiers journals use the numbered referencing system; to find the correct reference style and resources for the journal you are submitting to, please visit our help center.

Reference examples are found below, for more examples of citing other documents and general questions regarding the Vancouver reference style, please refer to Citing Medicine.

In-text citations

- Please apply the Vancouver system for in-text citations
- In-text citations should be numbered consecutively in order of appearance in the text – identified by Arabic numerals in the parenthesis (use square brackets for physics and mathematics articles).

Reference examples

Article in a print journal Sondheimer N, Lindquist S. Rnq1: an epigenetic modifier of protein function in yeast. *Mol Cell* (2000) 5:163-72.

Article in an online journal Tahimic CGT, Wang Y, Bikle DD. Anabolic effects of IGF-1 signaling on the skeleton. *Front Endocrinol* (2013) 4:6. doi: 10.3389/fendo.2013.00006

Article or chapter in a book Sorenson PW, Caprio JC. "Chemoreception". In: Evans DH, editor. *The Physiology of Fishes*. Boca Raton, FL: CRC Press (1998). p. 375-405.

Book Cowan WM, Jessell TM, Zipursky SL. *Molecular and Cellular Approaches to Neural Development*. New York: Oxford University Press (1997). 345 p.

Abstract Christensen S, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, editor. *Genetic Programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3–5; Kinsdale, Ireland*. Berlin: Springer (2002). p. 182–91.

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

Patent Pagedas AC, inventor; Ancel Surgical R&D Inc., assignee. Flexible Endoscopic Grasping and Cutting Device and Positioning Tool Assembly. United States patent US 20020103498 (2002).

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>

Theses and dissertations

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

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