

**MOLECULAR GENETICS: STRATEGIES TO IDENTIFY CONGENITAL  
CATARACT GENES IN CAPTIVE-BRED VERVET MONKEYS**

By

Zandisiwe Emilia Magwebu

A thesis submitted in partial fulfilment of the requirements  
for the degree of

**MAGISTER SCIENTIAE**

In the Department of Medical Biosciences

The University of the Western Cape

26 February 2013

Supervisors:

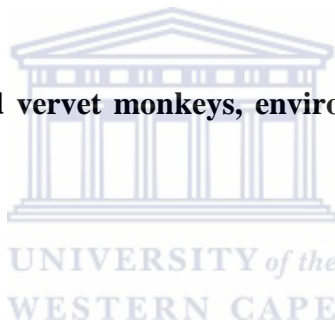
Dr: Jurgen Seier and

Dr Sahar Abdul- Rasool

## TABLE OF CONTENTS

<b>KEYWORDS</b>	<b>i</b>
<b>ABSTRACT</b>	<b>ii</b>
<b>DECLARATION</b>	<b>iv</b>
<b>LIST OF ABBREVIATIONS</b>	<b>v</b>
<b>LIST OF FIGURES</b>	<b>ix</b>
<b>LIST OF TABLES</b>	<b>xviii</b>
<b>ACKNOWLEDGEMENTS</b>	<b>xix</b>
<b>CHAPTER ONE: Literature review</b>	<b>1</b>
<b>1.1 Introduction</b>	<b>1</b>
<b>1.1.2 Cataract development in nonhuman primates</b>	<b>5</b>
<b>1.2 Phenotypic characterization of cataracts</b>	<b>6</b>
<b>1.2.1 Anterior polar cataracts</b>	<b>8</b>
<b>1.2.2 Posterior polar cataracts</b>	<b>8</b>
<b>1.2.3 Nuclear cataracts</b>	<b>9</b>
<b>1.2.4 Lamellar cataracts</b>	<b>9</b>
<b>1.2.5 Pulverulent cataracts</b>	<b>10</b>
<b>1.2.6 Aceuliform cataracts</b>	<b>10</b>
<b>1.2.7 Cerulean cataracts</b>	<b>10</b>
<b>1.2.8 Total cataracts</b>	<b>11</b>
<b>1.2.9 Cortical cataracts</b>	<b>11</b>
<b>1.2.10 Sutural cataracts</b>	<b>12</b>

<b>1.3 Congenital cataract genes</b>	<b>12</b>
<b>1.3.1 The Heat Shock transcription Factor 4 (<i>HSF4</i>) gene</b>	<b>14</b>
<b>1.3.2 Crystallin Alpha A (<i>CRYAA</i>) gene</b>	<b>15</b>
<b>1.3.3 The Glucosaminyl (N-acetyl) transferase 2 gene</b>	<b>17</b>
<b>1.3.4 The Lens Intrinsic Membrane protein 2 (<i>LIM2</i>) gene</b>	<b>18</b>
<b>1.4 The objectives of the study</b>	<b>20</b>
<b>1.5 Hypothesis</b>	<b>20</b>
<b>CHAPTER TWO: Materials and Methods</b>	<b>21</b>
<b>2.1 Ethical clearance</b>	<b>21</b>
<b>2.2 Management of captive bred vervet monkeys, environment and diet</b>	<b>21</b>
<b>2.3 Selection criteria</b>	<b>22</b>
<b>2.4 Blood collection</b>	<b>22</b>
<b>2.5 Extraction and purification of DNA from whole blood</b>	<b>23</b>
<b>2.6 Spectrophotometric quantification of nucleic acids</b>	<b>23</b>
<b>2.7 Bioinformatic identification of candidate genes and sequence variants</b>	<b>24</b>
<b>2.8 Standard Polymerase Chain Reaction (PCR)</b>	<b>25</b>
<b>2.9 Electrophoresis</b>	<b>26</b>
<b>2.9.1 Agarose gel</b>	<b>26</b>
<b>2.10 Purification of PCR product</b>	<b>27</b>
<b>2.11 Sequencing reactions</b>	<b>28</b>



<b>2.12 DNA sequence analysis</b>	<b>29</b>
<b>CHAPTER THREE: Results</b>	<b>30</b>
<b>3.1 Family trees of selected Vervet monkeys</b>	<b>30</b>
<b>3.2 DNA sequence variations in captive bred vervet monkeys</b>	<b>37</b>
<b>3.2.1 DNA Sequence variants in <i>HSF4</i></b>	<b>37</b>
<b>3.2.2 DNA sequence variants in <i>CRYAA</i></b>	<b>41</b>
<b>3.2.3 DNA sequence variants in <i>GCNT2</i></b>	<b>44</b>
<b>3.2.4 DNA sequence variants in <i>LIM2</i></b>	<b>58</b>
<b>CHAPTER FOUR: Discussion and Conclusion</b>	<b>59</b>
<b>4.1 Introduction</b>	<b>59</b>
<b>4.2 Limitations of the study</b>	<b>65</b>
<b>4.3 Advantages of the study</b>	<b>65</b>
<b>4.4 Conclusion</b>	<b>66</b>
<b>REFERENCES</b>	<b>68</b>
<b>APPENDIX A</b>	<b>82</b>
<b>APPENDIX B</b>	<b>90</b>





## KEYWORDS

Vervet monkeys

Congenital cataract

Inheritance

Autosomal Recessive

Cataract subtypes

PCR

*HSF4*

*CRYAA*

*GCNT2*

*LIM2*



## ABSTRACT

Molecular genetics: strategies to indentify congenital cataract genes in captive-bred Vervet monkeys

Zandisiwe Emilia Magwebu

MSc thesis, Department of Medical Biosciences, University of the Western Cape

The present study describes molecular aspects of inherited congenital cataract in captive-bred Vervet monkeys. Congenital cataracts are lens opacities that are present at birth or soon after birth and include hereditary cataracts or cataracts caused by infectious agents. The MRC Primate Unit is housing a colony of captive-bred Vervet monkeys in which 7.5% is suffering from congenital cataract. However, the parents of the affected individuals were asymptomatic. Six families within the colony have been identified to be affected by two types of morphologies (Y-sutural and total cataract). Based on the evidence provided above, it was speculated that the colony was affected with autosomal recessive cataract.

The main aim of this study was to facilitate a strategy for managing breeding programs by minimizing cataract occurrences in captive-bred Vervet monkeys. Integrated combination of clinical, molecular and bioinformatic strategies were used to identify and assess reciprocal candidate susceptibility genes for cataracts. The genes that are known to be responsible for most human congenital cataract cases were prioritized. The genes include Heat shock transcription factor 4 (*HSF4*), Crystalline Alpha A (*CRYAA*), glucosaminyl (N-acetyl) transferase 2 (*GCNT2*)

and Lens intrinsic membrane protein 2 (*LIM2*). Twenty two subjects were selected based on their morphology (5 carriers, 5 controls and 12 cataracts). 2ml of blood was collected for Deoxyribonucleic acid (DNA) extraction. Coding exons and flanking regions were screened by polymerase chain reaction (PCR) amplification and sequenced. The CLC DNA workbench was used for results analysis.

The screening of four genes revealed 20 sequence variants which were not present in the control individuals. Sequencing of *HSF4* revealed three mutations: R116R, L245>L and P421>L in exon 5, 10 and 14, respectively. The coding exons for *CRYAA* showed two sequence variants: S134W and K166N in exon 3. Twelve mutations were identified in exon one of all three *GCNT2* transcripts (A, B and C). These mutations include: G212G, H256>H, M258>V, N275>N, V16>I, Y122>F, S15>S, S24>N, S38>S, I118>I, D194>D and Y373>Y which was found in exon three of all transcripts. There were no mutations in *LIM2*, however, three single nucleotide polymorphisms (SNPs) were identified in exon 2 (P66>P) and 3 (I118>T and A127>T). The above mutations were conserved when aligned with other species. The sequence variations vary among the families and those individuals with the same or different cataract phenotype.

Based on these findings, it can be concluded that the four candidate genes harbour mutations that are responsible for both phenotypes. The effect of these mutations in Vervet monkeys is not yet understood, however, their impact will be further investigated. For future studies, it will be of absolute importance to screen the entire family to verify that indeed cataract formation in this colony is inherited in an autosomal recessive manner.

## DECLARATION

I, Zandisiwe Emilia Magwebu, declare that “**Molecular genetics: strategies to identify congenital cataract genes in captive-bred vervet monkeys**” is my own work and has not been submitted to any University. All the sources that I have used or quoted have been indicated and acknowledged by means of complete references.



Zandisiwe Emilia Magwebu

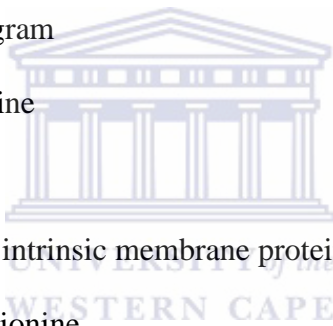
Date.....

Signed:

## LIST OF ABBREVIATIONS

A	Adenine
AA	Amino acids
ACD	Alpha crystallin domain
AD	Autosomal dominant
AR	Autosomal recessive
bp	Base pair
C	Celsius
CRYAA	Crystallin alpha A
D	Aspartate
d	Dye
dd	Double distilled
ddd	Triple distilled
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetate
<i>et al</i>	<i>et alia</i> (and others)
EtBr	Ethidium bromide
F	Phenylalanine
G	Guanine
G	Glycine
g	Grams
GCNT2	Glucosaminyl (N-acetyl) transferase 2

H	Histidine
H <sub>2</sub> O	Water
HCl	Hydrochloric acid
HR	Hydrophobic repeats
HSE	Heat shock elements
HSF4	Heat shock transcription factor 4
Hsps	Heat shock proteins
I	Isoleucine
K	Lysine
Kg	Kilogram
L	Leucine
L	Liter
LIM2	Lens intrinsic membrane protein 2
M	Methionine
MRC	Medical Research Council
min	Minutes
ml	Milliliter
mg	Microgram
mm	Millimeter
mM	Milimolar
N	Asparagine
ng	nanogram
nm	Nanometer



N/A	Not applicable
OD	Optical density
P	Proline
PCR	Polymerase chain reaction
pmol	Picomolar
R	Arginine
RBCs	Red blood cells
RNA	Ribonucleic acid
S	Serine
SNP	Single nucleotide polymorphisms
T	Thymine
Taq	<i>Thermus aquaticus</i>
TBE	Tris Borate EDTA
T <sub>m</sub>	Melting temperature
μl	Microliter
μg	Microgram
μM	Micromolar
UV	Ultraviolet
V	Volt
V	Valine
W	Tryptophan
WBC	White blood cells
Y	Tyrosine

%	Percentage
$\alpha$	Alpha
$\beta$	Beta
$\gamma$	Gamma

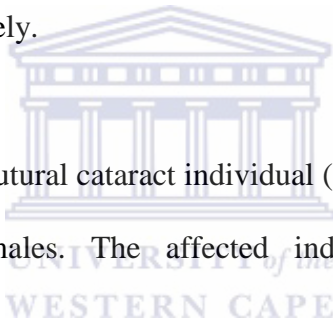




## LIST OF FIGURES

Figures		Pages
1.1	A schematic diagram illustrating normal and cataract crystalline lens (National Eye Institute, 2010)	2
1.2	Non-human primates with cataracts. <b>A)</b> New World monkey presented with bilateral cataracts (adapted from Deseret news, 2008). <b>B)</b> Old World monkey presented with unilateral cataracts (MRC Primate Unit).	5
1.3	Types of cataracts. <b>A:</b> Anterior polar cataract. <b>B:</b> Posterior polar cataract. <b>C:</b> Nuclear cataract. <b>D:</b> Lamellar cataract. <b>E:</b> Pulverulent cataract. <b>F:</b> Aceuliform-like cataract. <b>G:</b> Cerulean cataract. <b>H:</b> Total cataract. <b>I:</b> Cortical cataract. <b>J:</b> Sutural cataract. Adapted from Reddy <i>et al.</i> , (2004).	7
1.4	The chromosome structure of <i>HSF4</i> gene (GeneCards, 2012).	15
1.5	The chromosome structure of <i>CRYAA</i> gene (GeneCards, 2012)	17
1.6	The chromosome structure of <i>GCNT2</i> gene (GeneCards, 2012)	18
1.7	The chromosome structure of <i>LIM2</i> gene (Genecards, 2012)	19
2.1	100bp DNA ladder (Promega)	27

- 3.1** Pedigree B with two half-brothers with total cataract (371 and 398). The squares represent males and circles females. The affected individuals are marked (+) respectively. **30**
- 3.2** Pedigree A with four total cataract individuals (400, 409, 387, and 397) and one (311) Y-sutural cataract. The squares represent males, circles females, deceased with strike through (/) and the symbol E for individuals that were pair-bred in the E-cage where the parent is unknown. The affected individuals are marked (+) respectively. **31**
- 3.3** Pedigree C with one Y-sutural cataract individual (374). The squares represent males and circles females. The affected individuals are marked (+) respectively. **32**
- 3.4** Pedigree D with one total cataract individual (394). The squares represent males and circles females. The affected individuals are marked (+) respectively. **33**
- 3.5** Pedigree E with one total cataract individual (402). The squares represent males, circles females, deceased with strike through (/) and the E symbol: individuals that were pair-bred in the E-cage where the father was unknown. The affected individuals are marked (+) respectively. **34**



- 3.6** Pedigree F with two total cataract half-brother and sister (389 and 416). The squares represent males and circles females. The affected individuals are marked (+) respectively. **35**
- 3.7** **A)** One of the two monkeys with Y-sutural cataract. **B)** One of the ten monkeys with total cataract. The above morphologies were unilateral or bilateral in some individuals. **36**
- 3.8** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *HSF4* gene with one mutation in exon 5 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1313 C>T) found in exon 5 *HSF4*. The change at codon 116 (R116>R) was found in eight cataract and three carriers. **C)** The protein sequence alignment among different species where cataract represents for Vervet. **38**
- 3.9** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *HSF4* gene with one mutation in exon 10 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1700 C>T) found in exon 10 *HSF4*. The change at codon 245 (L245L) was found in one affected individual. **C)** The protein sequence alignment among different species where mutant represents Vervet. **39**

- 3.10** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *HSF4* gene with sequence variant in exon 14 (red circle). **B)** Sequence chromatogram showing transition missense mutation (c.5900C>T) found in exon 14 *HSF4*. The change at codon 421 was found in four cataract and two carriers, this resulted in the substitution of Proline to Leucine (P421L). **C)** The protein sequence alignment among different species where mutant represents Vervet. **40**
- 3.11** The structure of *HSF4* with previously reported mutations. Adapted from Shi *et al.*, 2008. The mutations that were identified in Vervet monkeys are indicated with green. **41**
- 3.12** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *CRYAA* gene with one mutation in exon 3 (red circle). **B)** Sequence chromatogram showing transversion missense mutation (c.470 C> G) found in exon 3 of *CRYAA*. The change at codon 134 was found in one carrier and five affected individuals, this resulted in substitution of Serine to Tryptophan (S134>W). **C)** The protein sequence alignment among different species where cataract represents Vervet. **42**
- 3.13** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *CRYAA* gene with one mutation **43**

in exon 3 (red circle). **B)** Sequence chromatogram showing transversion missense mutation (c.567 G > C) found in exon 3 of *CRYAA*. The change at codon 166 was found in three affected individuals, this resulted in substitution of Lysine to Asparagine (K166>N). **C)** The protein sequence alignment among different species where cataract represents Vervet.

**3.14** The structure of alpha crystalline genes with previously reported mutations. **44**

Adapted from Hansen *et al.*, 2007. The mutations that were identified in Vervet monkeys are indicated in green.

**3.15** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose **45**

gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.917G>A) found in exon 1 of *GCNT2* transcript A. The change at codon 212 (G212>G) was found in three carriers and eight affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

**3.16** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose **46**

gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1049 T>C) found in exon 1 of *GCNT2* transcript A. The change at codon 256 (H256>H) was found in one carrier and three affected individuals.

C) The protein sequence alignment among different species where mutant represents Vervet.

**3.17** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition missense mutation (c.1053A>G) found in exon 1 of *GCNT2* transcript A. The change at codon 258 (M258>V) was found in one carrier and three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **47**

**3.18** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1106 C>T) found in exon 1 of *GCNT2* transcript A. The change at codon 275 (N275>N) was found in three carriers and six affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **48**

**3.19** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition missense mutation (c.754 G>A) found in exon 1 of *GCNT2* transcript B. The change at **49**

codon 16 (V16>I) was found in one carrier and four affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

- 3.20** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transversion missense mutation (c.1073 A>T) found in exon 1 of *GCNT2* transcript B. The change at codon 122 (Y122>F) was found in two carriers and three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **50**
- 3.21** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.275T>C) found in exon 1 of *GCNT2* transcript C. The change at codon 15 (S15>S) was found in two affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **51**
- 3.22** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition missense mutation (c.301G>A) found in exon 1 of *GCNT2* transcript C. The change at **52**

codon 24 was found in one affected individual, this resulted in the substitution of Serine to Asparagine (S24>N). **C)** The protein sequence alignment among different species where mutant represents Vervet.

**3.23** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.344T>C) found in exon 1 of *GCNT2* transcript C. The change at codon 38 (S38>S) was found in three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **53**

**3.24** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.794 C> T) found in exon 1 of *GCNT2* transcript C. The change at codon 188 (I188>I) was found in five affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **54**

**3.25** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.812 C> T) found in exon 1 of *GCNT2* transcript C. The change at codon 194 (D194>D) was found in seven affected individuals. **C)** The protein **55**



sequence alignment among different species where mutant represents Vervet.

**3.26** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 3 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1670C> T) found in exon 3 of *GCNT2*. The change at codon 373 (Y373>Y) was found in six affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet. **56**

**3.27** The structure of *GCNT2* with previously reported mutations. Adapted from Pras *et al.*, 2004. The mutations that were identified in Vervet monkeys are indicated in green. **57**

**3.28** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *LIM2* gene with SNPs in exon 2 and 3 (red circles). **B)** The schematic structure of *LIM2* illustrating SNPs that were identified in Vervet monkeys (green) (Adapted from Shiels *et al.*, 2007). **C)** The protein sequence alignment among different species for exon 2 and 3, respectively. **58**

## LIST OF TABLES

<b>Tables</b>		<b>Pages</b>
<b>1.1</b>	Genes associated with congenital cataract and their mutations	<b>13</b>
<b>2.1</b>	Sample population	<b>22</b>
<b>2.2</b>	Sequencing reactions, as described in the Perkin-Elmer ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit manual (1998)	<b>28</b>
<b>2.3</b>	Sequencing reaction using a Hybaid PCR Sprint Thermal Cycler	<b>29</b>



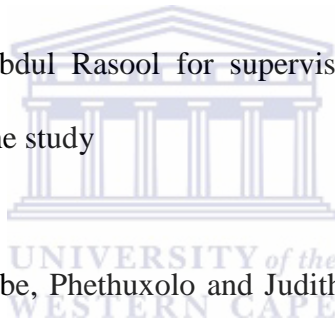
## ACKNOWLEDGEMENTS

**I would like to sincerely thank the following;**

Medical Research Council for financial support.

My mentor Dr CG Chauke for allowing me to make mistakes and learn from them and all the effort you have put to make this research a possible one, without your guidance I would not have accomplished this thesis.

Dr Jurgen Seier and Dr Sahar-Abdul Rasool for supervision, your willingness to help, the support and guidance throughout the study



My lab mates Joritha, Timothy, Abe, Phethuxolo and Judith who assisted me in collecting the samples.

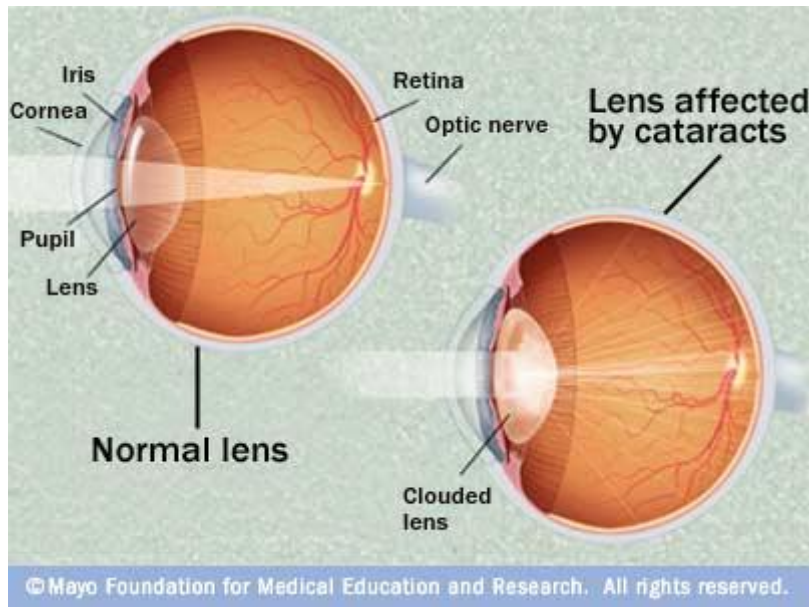
To my loving family for their love and support.

## CHAPTER ONE

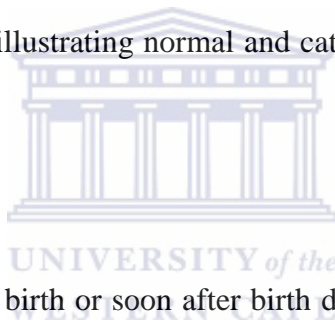
### Literature review

#### 1.1 Introduction

A cataract is a pathological condition that results in clouding of the normally clear crystalline lens of the eye (Song *et al.*, 1997; Francis *et al.*, 1999; Hejtmancik, 2008). The lens is located behind the pupil and allows light to pass through and be focused onto the retina to enable vision. Cataracts affect the normal function of the lens by interfering with the sharp focus of light in the retina (Figure 1.1). It is the most common treatable causes of visual loss in humans, and may be broadly divided into early onset (congenital or juvenile) (Ke *et al.*, 2006) and age related (Lambert and Drack, 1996; Foster, 1999). In South Africa there are approximately 330,000 blind people, 80% of whom live in rural areas. The Ophthalmological Society of South Africa (2007) estimates that there are more than 21 000 blind children in South Africa, and according to global statistics, 50% of childhood blindness is avoidable. Genetic factors are often a cause of congenital cataracts, and positive family history may play a role in predisposing someone to cataracts at an earlier age.



**Figure 1.1:** A schematic diagram illustrating normal and cataract crystalline lens (National Eye Institute, 2010)



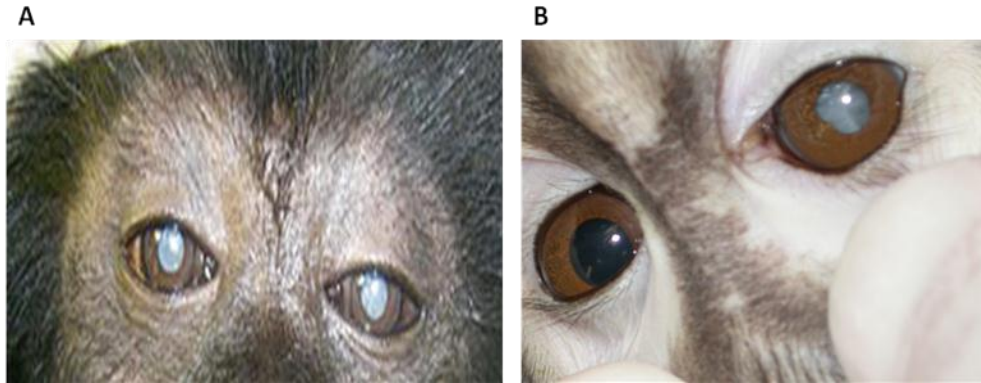
Congenital cataracts are present at birth or soon after birth due to an abnormal formation of the lens while in the womb (Ke *et al.*, 2006). These are either hereditary or caused by infectious agents and can occur in isolation or as a component of a syndrome affecting multiple tissues (Chen *et al.*, 2011). Similarly, a mutation causing a severe insult to the lens cell that result in major and immediate disruption of cell homeostasis can cause congenital cataract which is inherited in Mendelian fashion (Hejtmancik, 2008). It has the potential for inhibiting visual development, resulting in permanent blindness by interfering with the sharp focus of light on the retina (Graw *et al.*, 2001). In some cases, the inherited cataract is not significant enough to affect vision, if it does; however, it should be removed.

Cataracts can be inherited in an autosomal dominant, recessive or X-linked mode (Kannabiran and Balasubramanian, 2000). According to Hejtmancik (2008), phenotypically identical forms can result from mutations at different genetic loci, and they may have different inheritance patterns. Recently, many of the genes causing congenital cataracts have been identified, and have been associated with various developmental and forms of cataract although the number is constantly increasing (Zhang Q *et al.*, 2004; Hejtmancik 2008). It seems likely that when mutations in crystallin or other lens proteins are sufficient to cause protein aggregation, they usually result in congenital cataracts. However, if these genes increase susceptibility to environmental insults, such as light, hyperglycemic, or oxidative damage, they might contribute to age-related cataract (Hejtmancik & Smaoui, 2003). Thus, congenital cataracts tend to be inherited in a Mendelian fashion with high penetrance, while age-related cataracts tend to be multifactorial, with both multiple genes and environmental factors influencing the phenotype. This makes the latter significantly less conducive to genetic and biochemical studies (Shiels and Hejtmancik, 2007).

The molecular mechanisms that govern the cataractogenesis in humans are poorly understood (Lambert and Drack, 1996). Family histories are difficult to obtain, and this makes it challenging when undertaking genetic studies in humans (Zhang *et al.*, 1991). Several animal models with inherited cataracts have been used extensively in order to understand the disease processes better (Duy *et al.*, 2010). They readily provide tissue samples throughout the cause of the disease and are excellent model systems to investigate the underlying causes at developmental, morphologic, and biochemical level. They also enable the researcher to manipulate the mating to determine

whether an individual animal is pure bred for a specific trait. Among these animal models, rats and mice have received more attention during the last several years (Wolf *et al.*, 2000). Although the studies on rats and mice have provided important information on the biochemical and morphologic changes that occur during cataract development, investigations on other animal models such as nonhuman primates would be most informative (Zhang *et al.*, 1991).

Nonhuman primates are closely related to humans due to the fact that their genetic make-up is similar (King *et al.*, 1988; Vandenberg and Williams-Blangero, 1997). They share many specific genetic mechanisms which are involved in determining differential susceptibility to diseases. Anatomy and physiology of the human eye appears to be similar to primates even though the human lens differs in important ways from those of the most usually studied animals (rats and mouse) (Eichler and deJong, 2002). However, while the numerical differences in nucleic acids may be small, differences in gene and gene control mechanisms are prominent and form the bases for the phenotypic distinctions between humans and primates (King *et al.*, 1988). In many basic and applied studies, primates are the only appropriate animal model when they are only susceptible to the disease under study (Rogers *et al.*, 2006), or when they possess the biological or behavioural characteristics needed to investigate the scientific question most effectively (Vandenberg and Williams-Blangero, 1997). Therefore, the high degree of genetic similarity between humans and nonhuman primates result in physiological similarities and this makes the monkey an important resource for biological studies of humans (Vandenberg, 1995. For many years cataracts were mainly investigated in humans, while Old and New World monkeys were not given attention even though they also develop cataracts (Figure 1.2).



**Figure 1.2:** Non-human primates with cataracts. **A)** New World monkey presented with bilateral cataracts (adapted from MacLaren, 2008). **B)** Old World monkey presented with unilateral cataracts (MRC Primate Unit).

### 1.1.2 Cataract development in nonhuman primates

Cataracts have been reported in a colony of captive-bred Vervet monkey (*Chlorocebus aethiops*) colony (de Villiers *et al*, 2001). The onsets of the cases were observed at an age of between 3 months and 1 year and it is tempting to speculate that the cataracts in our Vervet monkeys are congenital, since they could have been present at birth microscopically (de Villiers *et al.*, 2001). Over the years, it became obvious that this was a colony-wide systematic problem. The clinical presentation was characterized by total and Y-sutural cataracts. These phenotypes suggested that the colony might be affected by different types of the disease and since available parents of affected individuals were asymptomatic, a possibility of autosomal recessive transmission mode exists.



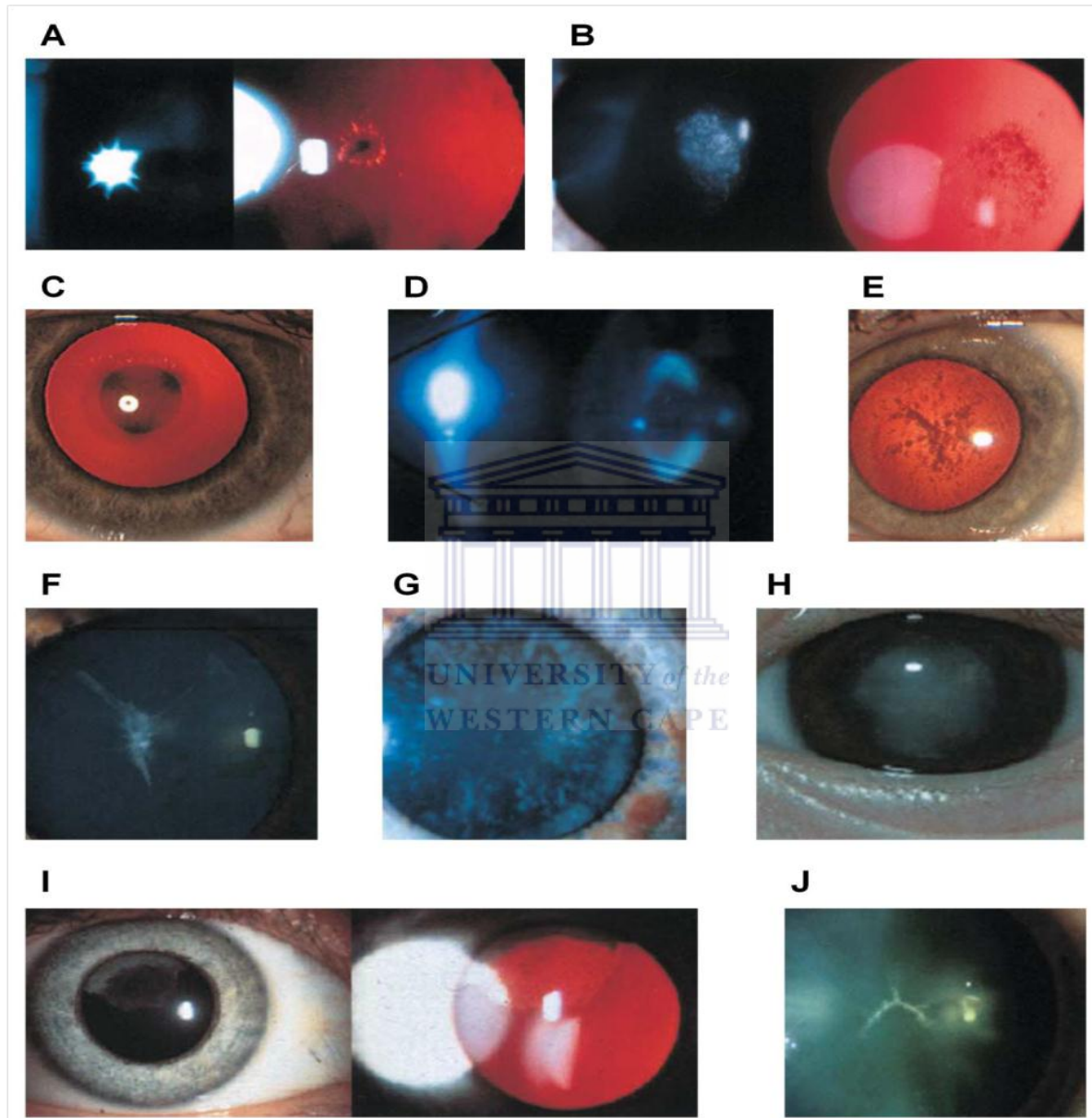
Cataracts occur spontaneously in a number of non-human primates (Ollivet *et al.*, 2000) whereas it can be induced in laboratory animals especially rodents. So far, no cases of cataracts have been associated with infectious agent (Plesker *et al.*, 2005). Cataracts have been reported in Vervet monkeys (Plesker *et al.*, 2005; de Villiers *et al.*, 2001; Souri, 1973), Woolly monkey (Peiffer and Gelatt, 1976), Rhesus macaque (Kessler and Rawlins, 1985) and Cynomolgus monkey (Sasaki *et al.*, 2011) and most of these cases were case studies. All reports are from animals in captivity at the time of diagnosis, and the disease was speculated to be of genetic origin. Captive breeding is often conducted with a small gene pool, and with inbreeding resulting in genetic similarity.

## 1.2 Phenotypic characterization of cataracts

Currently available literature review is mainly about humans since there is very limited data on congenital cataract in Vervet or Rhesus monkeys.

Cataracts are classified according to their morphological appearance, size, density, anatomic location and progression of opacity. However, making the diagnosis of a specific type of cataract can be difficult if it spreads to involve multiple layers, obscuring the original opacity (Ionides *et al.*, 1999). Over the years, different types of cataract have been identified and can be observed in either one or both eyes affecting adults and infants (Reddy *et al.*, 2004). Congenital cataracts can be classified into several subtypes according to morphology: anterior, posterior polar, nuclear, lamellar, pulverulent, aceuliform, cerulean, total, cortical and sutural (Yamada *et al.*, 2000; Shiels and Hejtmancik, 2007). These phenotypes described below (Figure 1.3) vary in age of

onset, severity and morphology in different families and even within the same family (Smaoui *et al.*, 2004).



**Figure 1.3:** Types of cataracts. **A:** Anterior polar cataract. **B:** Posterior polar cataract. **C:** Nuclear cataract. **D:** Lamellar cataract. **E:** Pulverulent cataract. **F:** Aceuliform-like cataract. **G:** Cerulean cataract. **H:** Total cataract. **I:** Cortical cataract. **J:** Sutural cataract. Adapted from Reddy *et al.*, (2004).

### **1.2.1 Anterior polar cataracts**

Anterior polar cataracts are a clinically well-recognized opacity located at the front of the lens and represents 3–14% of all congenital cataracts. It is situated at the anterior pole of the lens and is usually symmetrical and discrete lesions (Figure 1.3A) (Reddy *et al.*, 2004). Clinical presentation varies between different family members and it is unilateral most of the time, however, if bilateral they usually asymmetrical Patients with bilateral anterior polar cataracts are at high risk of developing amblyopia (Lambert and Drack, 1996)

### **1.2.2 Posterior polar cataracts**

A posterior polar cataract is a round, discoid, opaque mass that is composed of malformed and distorted lens fibers located in the central posterior capsule of the lens (Figure 1.3B) and consists of characteristic circular plaque with concentric whorls (Amaya *et al.*, 2003). The opacity extends anteriorly into the posterior cortex and is thicker than the standard posterior subcapsular cataract. It is a common type of visually symptomatic congenital cataract. Although sporadic cases of this type of cataract exist, it is usually inherited in an autosomal-dominant manner. Dominantly inherited posterior polar congenital cataract has previously been linked with the haptoglobin locus (*HP*) on 16q, along with the dominant Marner congenital cataract (Ionides *et al.*, 1999).

### 1.2.3 Nuclear cataracts

A nuclear cataract refers to opacification of the central zone of the lens, specifically the region between the anterior and posterior Y sutures (Lambert and Drack, 1996). The cataract is referred to as riders due to cortical fibers encircling the nuclear opacity and is located within the fetal nuclei of the lens (Figure 1.3C) (Hejtmancik, 2008). It can be inherited as autosomal dominant, recessive or X-linked. This type of cataract causes a gradual yellow cloudiness and hardening of the central part of the lens called the nucleus. Changes in vision are usually gradual and suggest an abnormality of gene expression in early development (Francis *et al.*, 2000). In some cases, patients show bilateral symmetrical involvement with variable expressivity.



### 1.2.4 Lamellar cataracts

Lamellar cataracts are characterized by a lamella of lenticular opacification sandwiched between a clear nucleus and cortex (Lambert and Drack, 1996). The opacity only affects some of the layers of the eye lens i.e. the layers outside the nucleus (Figure 1.3D). Lamellae are formed by a deposition of the secondary lens fibers that occurs during the growth of the normal lens (Rogaev *et al.*, 1996). This is the most common type of congenital cataract; it is inherited as autosomal dominant and usually seen bilaterally in early childhood (Bu *et al.*, 2002). The size of the opacity can give an indication as to when the cataract arose. In some cases, cortical cataract is associated with lamellar cataract.

### **1.2.5 Pulverulent cataracts**

It is characterised by a dust like appearance of the opacity which can be found in any part of the lens (Figure 1.3E) (Francis *et al.*, 2000). The embryonic and fetal nuclei are usually involved and diffuse cortical opacities may also be seen in some patients. The involved area is therefore larger than the somewhat similar to Coppock-like cataract which is limited to the embryonic nucleus (Gill *et al.*, 2000). This type of cataract has been documented in autosomal dominant and recessive disease (Rees *et al.*, 2000). Some patients presented pulverulent cataract restricted to nucleus or lamellae.



### **1.2.6 Aceuliform cataracts**

Aceuliform cataracts are associated with needle-like projections extending from the nucleus into the anterior and posterior cortex; there may well be overlap among the three phenotypes (Figure 1.3F) (Reddy *et al.*, 2004). According to Heon *et al.*, (1998) the needle-like crystals originate from the fetal and postnatal nuclei and project in different directions, through or close to the axial region of the lens.

### **1.2.7 Cerulean cataracts**

It is an unusual phenotype characterised by a tiny blue and white opacity which generally appears from birth through 18 and 24 months of age but may not be diagnosed until adulthood (Figure 1.3G) (Francis *et al.*, 2000). They first appear at the outer age of the fetal lens nucleus or

in more superficial cortical layers depending on the type. The opacities are usually bilateral and progressive. They consist of cerulean blue and white specks distributed across the lens in a scattered pattern, with a greater distribution density in the cortex. Lens removal may be required in early infancy. Cerulean cataracts of congenital or childhood onset can be due to mutations in genes that encode various lens crystallins (Berry *et al.*, 2001).

### **1.2.8 Total cataracts**

It is the opacity that affects both nuclear and cortical regions which according to Francis *et al.*, (2000) has been reported in families with autosomal dominant as well as X linked recessive congenital cataract (Reddy *et al.*, 2004). Total cataract may occur as a result of progression of other phenotypes of cataracts if left untreated. If an individual is affected with total cataract, the entire lens will be cloudy (Figure 1.3H) (Reddy *et al.*, 2004). They are frequently bilateral and begin as lamellar or nuclear cataracts (Lambert and Drack, 1996). This type of cataract was observed in ten of MRC Primate Unit captive bred-Vervet monkeys (De Villiers *et al.*, 2001).

### **1.2.9 Cortical cataracts**

Cataracts limited to the cortex are rare and the sector of the lens is affected in the outer cortex adjacent to the capsule (Reddy *et al.*, 2004). A cortical cataract generally appears as a cloudy opacity in the cortex (Figure 1.3I). These cataracts often resemble wheel spokes that point inward toward the center of the lens. Light tends to scatter when it hits the spoke-like opacities.

In the case of cortical cataract, the nucleus is not affected (Ionides *et al.*, 1999; Reddy *et al.*, 2004). However, the distribution and progression suggest an abnormality of the later stages of lens development even though the pathogen is unknown (Francis *et al.*, 2000).

### **1.2.10 Sutural cataracts**

The sutural also known as stellate is defined as an opacity affecting the whole or part of the anterior or posterior suture of one or both eyes of the fetal nucleus, at which the ends of the lens fiber cells converge (Figure 1.3J) (Hejtmancik, 2008). According to Vanitha *et al.*, (2001) the anterior and posterior Y sutures have prominent, dense white opacities. The shape and color may vary from patient to patient. However, it has been reported to be congenital without progression and have been described as being inherited as both autosomal dominant and X-linked traits (Klopp *et al.*, 2003). This type of cataract was observed in two of MRC Primate Unit captive bred-Vervet monkeys (De Villiers *et al.*, 2001).

### **1.3 Congenital cataract genes**

To date, at least 39 loci in the human genome have been reported to be associated with various forms of congenital and developmental cataracts (Zhang Q *et al.*, 2004). Among them, mutations in 24 genes responsible for such cataracts have been identified, most of which belong to seven groups; (1) crystallins, the most abundant proteins in the lens, (2) enzymes necessary for maintaining lens metabolism, (3) membrane proteins, (4) cytoskeletal proteins, (5) protein

participating in ion transport, (6) transcriptional factors and (7) genes with as yet undefined functions (Azuma *et al.*, 2000; Zhang Q *et al.*, 2004). Autosomal dominant congenital cataract appears to be the most common familial form, autosomal recessive and X-linked cataracts also occur.

**Table 1.1:** Genes associated with congenital cataract and their mutations

LOCUS	GENE	INHERITENCE	PHENOTYPE	DNA MUTATION	AMINO ACID
16q21-22.1	<i>HSF4</i>	AD	Total	R74H	c221G > A
		AD	Zonular	c348T>C	L115P
		AD	Zonular	c362C>T	C120R
		AR	Zonular	c524G>C	R175P
		AR	Zonular	595delGGGCC	199fs
		AR	No data	c1213C>T	R405X
		AR	Total	IVs1327+4A>G	
		AR	Zonular	IVs1327+4A>G	
		19q13.4	<i>LIM2</i>	AR	Total
6p24	<i>GCNT2</i>	AR	Zonular	c987G>A	W328X W326X
21q22.3	<i>CRYAA</i>	AR	No data	c27G>A	W9X
		AD	Zonular	c34C>T	R12C
		AD	Zonular	c61C>T	R21W
		AD	Zonular	c145C>T	R49C
		AR	Total	c160C>T	R54C
		AR	No data	c160C>T	R54C
		AD	Zonular	c160C>T	R54C
		AD	Zonular	c346C>T	R116C
		AD	Zonular	c337G>A	R116H
		AD	Nuclear	c346G>A	R116H
		AD	Total	c346G>A	R116H
		AD	Zonular	c414G>A	R116H

Adapted from Huang and He (2010)

AD= Autosomal dominant

AR= Autosomal recessive

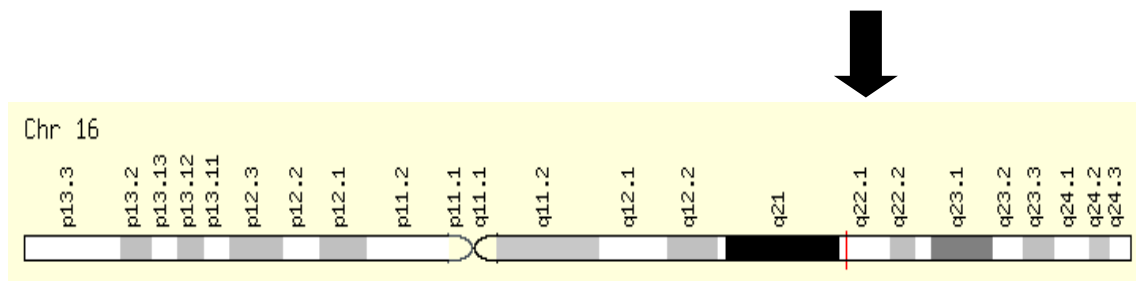


### 1.3.1 The Heat Shock transcription Factor 4 (*HSF4*) gene

Heat Shock transcription Factor 4 (*HSF4*) gene belongs to the heat shock transcription factor family. As shown in Figure 1.4, the human gene is located 16q22.1 (Sajjad *et al.*, 2008) whereas in Rhesus macaque it is located in chromosome 20, it has 15 exons (Shi *et al.*, 2008) and isoform *HSF4a* and *b* respectively (Smaoui *et al.*, 2004). The *HSF4a* represses *HSF4* genes by binding to heat shock elements (HSE), and the molecular mechanisms of HSF-4a isoform-mediated transcriptional repression is unknown, while *HSF4b* acts as activator of transcription (Frejtag *et al.*, 2001). The *HSF4* is known to regulate expression of heat shock proteins (Hsps) as a result of different stresses, such as oxidants, elevated temperature, heavy metals and bacterial and viral infections (Smaoui *et al.*, 2004). It has been reported for different modes of Mendelian inheritance such as autosomal recessive and dominant cataracts. The members of this family have three functional regions: DNA binding domain (DBD) located at the N-terminal half of *HSF4*, linker region and hydrophobic repeat regions (HR-A and B) with coiled-coil structures (figure 3.8) (Nakai *et al.*, 1997). Based on the structure, exons 3-5 and part of exon 6 encode part of DBD.

Sajjad *et al.*, (2008) reported the first nonsense mutation in Afghani/Pakistan family. The mutation Arg405X (Table 1.1) in exon 11 was predicted to cause a premature termination, resulting in a complete loss of function. In addition, two mutations (p.Arg175Pro and c.595\_599delGGGcc) were also reported in Pakistan family by Forsheew *et al.*, (2005) which were located in the HR-A and HR-B of *HSF4* (Table 1.1). It was observed that certain sequence changes of *HSF4* permit abnormal expression of Hsps. It is suspected that *HSF4* is a genetic

factor which is also associated with age-related cataractogenesis (Bagchi *et al.*, 2002). This gene is expressed in both cell types of the lens, such as epithelial and fiber cells (Fugimoto *et al.*, 2004) which require *HSF4* for normal cell growth and differentiation.



**Figure 1.4:** The chromosome structure of *HSF4* gene (GeneCards, 2012).

### 1.3.2 Crystallin Alpha A (*CRYAA*) gene



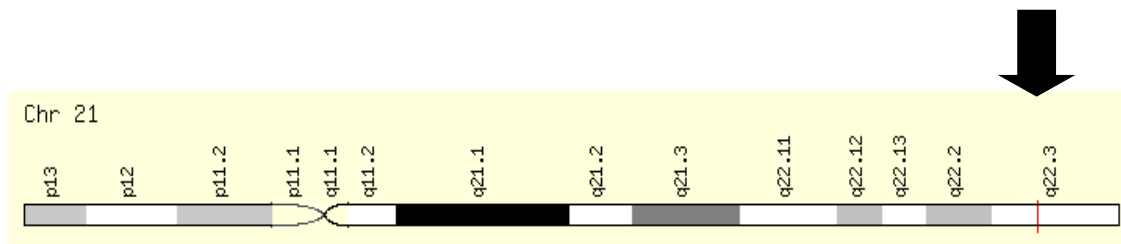
According to Andley (2007), crystallins are the predominant structural proteins in the lens that are evolutionarily related to stress proteins. Crystallins are known to constitute about 80-90% of water soluble proteins of the lens and contribute to the transparency and refractive proteins. Mammalian lens crystallins are divided into alpha, beta and gamma families, which are found in all vertebrate lenses and are referred as ubiquitous crystallins (Devi *et al.*, 2008). Alpha and beta families are further divided into acidic and basic groups (Hejtmancik, 1998). In the mature human lens,  $\alpha$ -crystallins makes up roughly 40%,  $\beta$ -crystallins 35% and  $\gamma$ -crystallins 25% of the total crystallin protein (Devi *et al.*, 2008). According to Kumar (1999), the  $\alpha$  crystallins belong to the small heatshock protein (sHSP) family which functions as molecular chaperones which are known to facilitate the correct folding of proteins *in vivo* and are important in keeping these proteins in a functional state. Thus, has the ability to protect the lens against oxidative stress

(Graw, 2009). There are two subunits;  $\alpha$ A and  $\alpha$ B, which are encoded by individual genes, *CRYAA* and *CRYAB* being localized on different chromosomes and have been characterised in various species (e.g. mouse, rat, chicken, rabbit, man and hamster) (Graw, 2009).

The *CRYAA* is more abundant in the eye and mutations that are expressed in the eye lens are expected to cause cataract only. It facilitates the correct folding of other lenticular proteins *in vivo*, thus suppressing protein aggregation and maintaining lens transparency (Khan *et al.*, 2007). It also defend against protein aggregation, act as molecular chaperones and have three functional regions; N-terminal,  $\alpha$  crystalline domain (ACD) and C-terminal which are used for higher assembly of dimmers (Horwitz, 1992). The ACD region consists of  $\beta$  sheets (B3-9) where B6 and 7 form a sandwich. The sheet is dynamic and can change shape in order to regulate a potential peptide binding sites. In addition, the dynamics of the  $\beta$  sandwich domain is essential for normal function of  $\alpha$  crystallines (Clark *et al.*, 2012).

In humans it is located in 21q22.3 (Figure 1.5), in Rhesus macaque on chromosome 3 and has 3 exons. The gene contributes to the structural and refractive properties and prevents oxidative damage (Piatigorsky, 1990). Mutations in the *CRYAA* lead to recessive or dominant cataracts (Graw, 2009). A W9X substitution in *CRYAA* in a Persian Jewish family caused a chain termination, and cataracts were inherited in an autosomal recessive manner (Pras *et al.*, 2000). Since the affected individuals had the cataracts extracted before three months of age, no cataract phenotype information was available for this mutation (Table 1.1). Moreover, further

experimental evidence suggest that  $\alpha$  crystallin is involved in remodelling and protection of the cytoskeleton, inhibition of apoptosis and the resistance to stress (Andley, 2007).

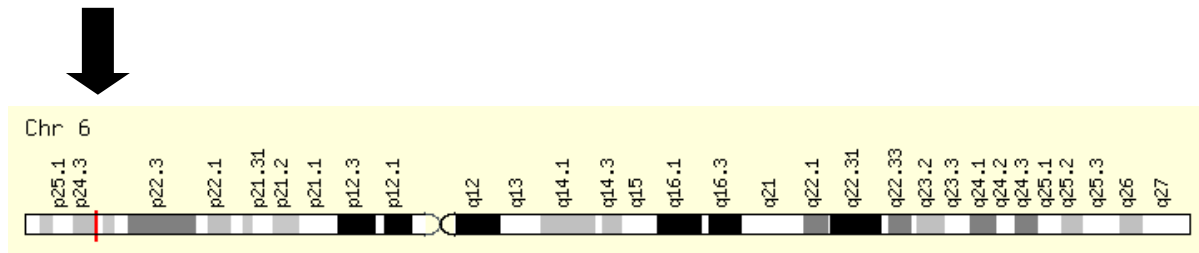


**Figure 1.5:** The chromosome structure of *CRYAA* gene (GeneCards, 2012)

### 1.3.3 The Glucosaminyl (N-acetyl) transferase 2 gene

It encodes glucosaminyl (N-acetyl) transferase 2, an enzyme that is responsible for formation of blood group I antigen. The i and I antigens are carbohydrate structures that are distinguished by linear and branched poly-N-acetylglucosaminoglycans, respectively (Song *et al.*, 1997). According to Pras *et al.*, (2004) the change from linear i carbohydrate to branched I structure takes place in human erythrocytes as *GCNT2* branching enzyme begins to be expressed. Adult human red blood cells fully express I antigens and contain only few fetal i antigens, however, some individuals have high levels of i antigen and this is referred to as adult i phenotype (Yu *et al.*, 2003). Even though the phenotype is rare, it has been linked to congenital cataract. In humans, *GCNT2* is located in chromosome 6p24 (Figure 1.6) whereas in the Rhesus macaque is in chromosome 4 and has three subunits (A, B & C) with identical exon 2 and 3 but differ in exon 1 (Pras *et al.*, 2004). The *GCNT2* isoforms are expressed in different cells for an example isoform C is expressed in erythrocytes and B form is expressed in the lens. A study done by Pras *et al.*,(2004) revealed a homozygous G3A substitution in base 58 of exon-2, resulting in the

formation of premature stop codons W328X, W326X, and W328X, of the *GCNT2A*, *-B*, and *-C* isoforms, respectively (Table 1.1). This mutation was associated with autosomal recessive congenital cataract in four distantly related Arab families from Israel.

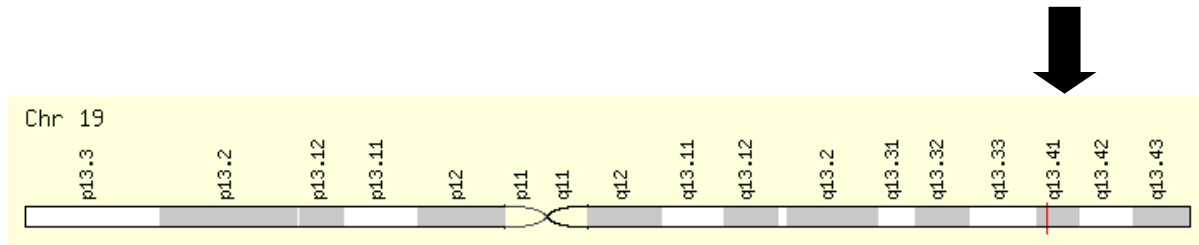


**Figure 1.6:** The chromosome structure of *GCNT2* gene (GeneCards, 2012)

### 1.3.4 The Lens Intrinsic Membrane protein 2 (*LIM2*) gene

The *LIM2* gene is the second most abundant intrinsic membrane protein in the lens fiber cells (Huang and He 2010). As shown in Figure 1.7, it is located in 19q13.4 in both humans and Rhesus macaque and has 5 exons which fall within the coding region of the sequence (Lieuallen *et al.*, 1994). Its function is not clearly understood as yet (Ponnam *et al.*, 2008). However, it encodes an eye lens-specific protein found at the junctions of lens fiber cells, where it may contribute to cell junctional organization (Tenbroek *et al.*, 1992). It has been shown to be absent from proliferating epithelial cells in the lens with expression becoming prominent in differentiating as well as in mature lens fiber cells (TenBroek *et al.*, 1994). It binds calmodulin (Louis, 1990) as well as galectin; a protein associated with lens cell membranes and may play an important role in both lens development and cataractogenesis (Gonen *et al.*, 2001). Mutations in this gene have been associated with cataract formation. Ponnam *et al.*, (2008) reported the

second known mutation (G154E) in *LIM2* causing autosomal recessive cataracts. The G154E mutation was found to be deleterious to protein function (Table 1.1).



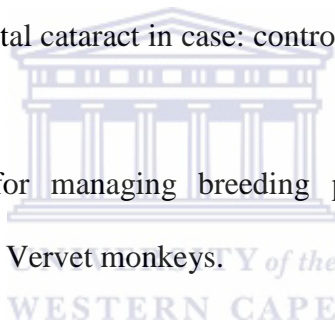
**Figure 1.7:** The chromosome structure of *LIM2* gene (Genecards, 2012)



## 1.4 The objectives of the study

To use an integrated combination of clinical, molecular biological and bioinformatic strategies to identify and assess reciprocal candidate susceptibility genes for cataracts:

1. To identify candidate susceptibility genes harboured in loci that have been reported to be linked to congenital cataract in humans, by using bioinformatic (computer-based) search tools.
2. To identify DNA sequence polymorphisms within newly recognized candidate genes by bioinformatic and molecular biological techniques and to assess the possible involvement of these genes with congenital cataract in case: control studies in a colony of captive-bred Vervet monkeys.
3. To facilitate a strategy for managing breeding programs by minimizing cataract occurrences in captive-bred Vervet monkeys.



## 1.5 Hypothesis

The Vervet monkey is a unique model for cataracts, which are inherited as an autosomal recessive trait and expressed as early-onset cataract in heterozygotes. The monkey eye is phylogenetically close to the human eye, and the occurrence of the cataracts in Vervet monkeys presents a valuable model in the field of cataractogenesis and primate genetics.

## CHAPTER TWO

### Materials and Methods

#### 2.1 Ethical clearance

The study was approved by Ethics Committee of the MRC (Ref 10/11). The project was carried out in the Primate Unit of the Technology and Innovation Directorate of the MRC.

Vervet monkeys are housed and maintained according to the National Guidelines for the Care and Use of Animals for Scientific Purposes.

#### 2.2 Management of captive bred vervet monkeys, environment and diet

The Vervet monkeys used in this study were bred and maintained under identical environmental conditions and received the same diet. The closed indoor environment was maintained at a temperature of 24-26°C, 40-70% humidity, 15-20 air changes/hour and a photoperiod of 12 hours (Seier *et al.*, 2012).

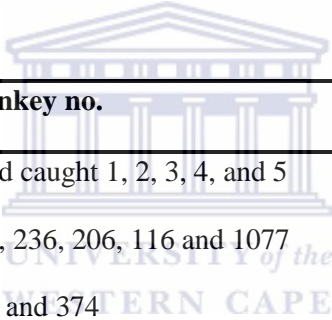
All diets were prepared in the Primate Unit and all primates were fed three times a day. The diet consisted of pre-cooked maize meal mixed with a vitamin and mineral concentrate, egg powder, bean flour and sunflower oil. Vitamin C and D3 were supplemented daily due to lack of exposure to sunlight. The monkeys received fruit during the day washed in chlorinated water. A mixture of ground and whole kernel maize was provided in the afternoon. Water was available via an automatic watering device.



### 2.3 Selection criteria

Selection of subjects included in this study was based on their family history and morphology. Animals that were considered controls are those that were not producing individuals with cataracts throughout their generation and are housed together with those that displayed the phenotype. The carriers were included as positive controls in order to verify the origin of the mutation. There is cataract in their ancestry especially the last generation which displayed cataract phenotypes (Cataract) (Table 2.1)

**Table 2.1:** Sample population.



<b>Subjects</b>	<b>Monkey no.</b>
<b>Controls*</b>	Wild caught 1, 2, 3, 4, and 5
<b>Carriers **</b>	215, 236, 206, 116 and 1077
<b>Cataracts: Y-sutural</b>	311 and 374
<b>Total</b>	371, 387, 389, 394, 397, 398, 400, 402, 409 and 416

\*Vervet monkeys from the wild

\*\*Parents of cataract individuals

### 2.4 Blood collection.

Blood (2ml) was obtained via femoral venipuncture after Ketamine anaesthesia at 5mg/kg bodyweight. Blood was collected in EDTA-containing tubes after overnight fasting. The blood samples were stored at -80°C.

## **2.5 Extraction and purification of DNA from whole blood.**

AxyPrep Blood Genomic Deoxyribonucleic acid (DNA) Miniprep Kit (Separations) was used to extract DNA from whole blood of Vervet monkeys. This method is based on the efficient release of genomic DNA from anti-coagulated whole blood by a special cell lysis and heme/protein precipitation buffer (Buffer AP1) coupled with the selective adsorption of the genomic DNA to a special AxyPrep column. The purified genomic DNA is eluted in a low-salt Tris buffer containing 0.5 mM EDTA, which enhances DNA solubility and helps to protect the high molecular weight DNA against subsequent nuclease degradation. Blood genomic DNA is directly isolated from the white blood cell (WBC) component of whole blood, without the need to remove the red blood cells (RBCs) in advance.

The procedure was performed according to the manufacturer's instructions (appendix A2). For maintaining the integrity and reactivity of the genomic DNA, particularly in PCR, the purified genomic DNA was eluted and stored in low-salt Tris buffer containing 0.5-1 mM EDTA.

## **2.6 Spectrophotometric quantification of nucleic acids**

DNA quantification was performed at the wavelengths of 260 nm and 280 nm. The reading at 260nm allows calculation of the concentration of nucleic acid in the sample, where an OD of 1 corresponds to approximately 50ug/mL for double-stranded DNA. The ratio between the readings at 260nm and 280 nm ( $OD_{260}/OD_{280}$ ) provides an estimate of the purity of the nucleic acid, with pure preparations of DNA giving  $OD_{260}/OD_{280}$  values of 1.8 (Maniatis, 1989).

Accurate quantification of the amount of nucleic acid is only possible if there is no protein or phenol contamination of the sample.

## **2.7 Bioinformatic identification of candidate genes and sequence variants**

Genes were located within the previously published cataract-linked loci by bioinformatic searches of public domain databases containing annotated genes (NCBI GENBANK), as well as by applying gene prediction programs (ENSEMBL). The UCSC genome browser (<http://genome.ucsc.edu>) was used to compare the Human sequence into Rhesus macaque and this assisted in identifying the chromosome number of the latter. The candidacy of genes selected by position was strengthened if they are expressed in the lens and/or show homology to genes encoding enzymes or receptors with roles in the crystallin architecture. Primers were designed using NCBI (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). The Rhesus macaque and Human sequences were used as a reference template since the Vervet monkey sequence is not yet published. The designed primers used in this study are in appendix A3 (Table A1).

Tools such as ClustalW and NCBI were used for cross-species sequence comparisons. Comparative analysis of DNA sequences from multiple species at varying evolutionary distances is a powerful approach for identifying coding and functional noncoding sequences, as well as sequences that are unique for a given organism. Comparing the DNA sequences of different species is a great method for decoding genomic information, because functional sequences tend to evolve at a slower rate than nonfunctional sequences. By comparing the genomic sequences of

species at different evolutionary distances, one can identify coding sequences and conserved noncoding sequences with regulatory functions, and determine which sequences are unique for a given species and may be responsible for traits that are unique to the reference species.

## 2.8 Standard Polymerase Chain Reaction (PCR)

Selected genes were amplified by polymerase chain reaction (PCR) which is an *in vitro* enzymatic amplification of defined DNA sequences to produce a high yield of amplified target DNA. This happens under the influence of specifically designed primers and a thermostable DNA polymerase of *Thermus aquaticus* (Taq). The process consists of three distinct steps. The first is denaturation of double stranded DNA followed by annealing of the primers to their complementary sequences on the template. The last step is extension by incorporation of nucleotides under the influence of Taq polymerase. With numerous repetitions of this set of steps, the number of copies of the target sequence rises exponentially.

In order to amplify target DNA, 0.2 ml thin walled tubes were used in a G-storm gradient thermocycler S/N 20024 (Vacutec) equipped with a heated lid. Unless otherwise stated, the standard 25  $\mu$ l-PCR reaction contained the following reagents: 2x PCR Master Mix (Promega), DNA template (50 ng) and 0.5  $\mu$ M of the upstream and downstream primers were made up to 25  $\mu$ l with Nuclease-Free water (appendix A4, Table A2).

The PCR conditions were adapted from Promega protocol and performed at : 95 °C for 5 min followed by 30 cycles of 95 °C for 30 sec, X °C for 30 sec and 72 °C for 1 min; an extension period of 5 min at 72 °C completing the procedure. X denotes the relevant annealing temperature which was chosen 5 °C below the assumed primer melting temperatures calculated using the following formula ( $T_m = [\text{no. of GC}] \times 4 + [\text{no. of AT}] \times 2$  °C. (appendix A4, Table A3). A DNA Sequencer (Applied Biosystems ABI3730xl DNA analyser) was used to obtain genotyping results

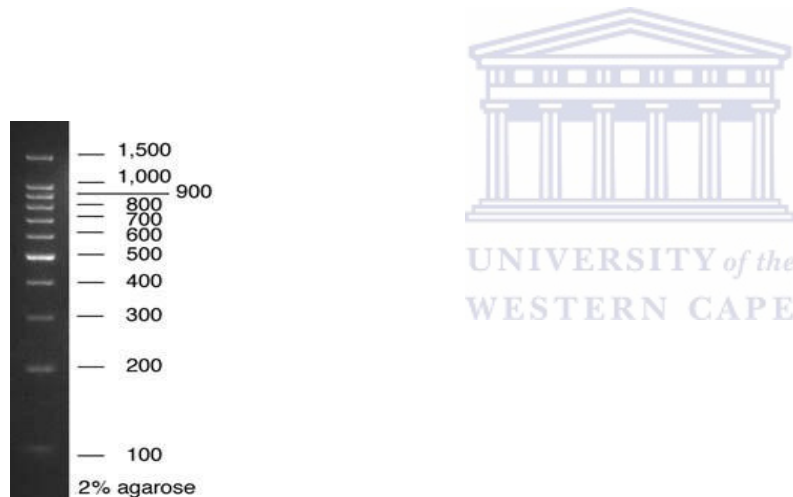
## **2.9 Electrophoresis**

Nucleic acids were separated by agarose gel electrophoresis technique based on charge, size and conformation. DNA is negatively charged and will migrate away from the negative pole through the agarose gel to the positive pole when the gel is placed in buffer in an electrical field. The molecules will separate out with the larger and more folded molecules moving slower through the gel and settling closer to the origin than the smaller less folded molecules that are able to move faster through the gel. Markers with proteins at known sizes are run on all gels as reference for determination of size of test DNA.

### **2.9.1 Agarose gel**

Agarose gels were made by dissolving the appropriate amount of agarose in 1X TBE buffer (appendix A5-A7) (12.1 g Tris, 0.37 g EDTA and 5.14 g Boric acid made up to 1L and adjusted to pH 8.4 with 1 M HCl) for 0.8 - 1.5 % gels, depending on the fragment size loaded onto the

gel. Genomic DNA was run on 0.8% gels, whereas 1.5% gels were used with fragment sizes of 190 bp and 550 bp. The agarose gels were electrophorized in TBE buffer at a voltage range between 80 - 120 V for approximately 1 hour. Samples were loaded into the wells with 10% tracking dye (appendix A8). 0.5mg/μl ethidium bromide (EtBr) added to the gels to allow visualisation of DNA when it was placed on a UV transilluminator, which caused any DNA bound to ethidium bromide to fluoresce. A 100bp DNA Ladder from Promega was used as a marker to identify the molecular weight of amplified exons (Figure 2.1). This Ladder is ready for 5' end-labeling with radioisotopes using T4 Polynucleotide Kinase, allowing visualization by autoradiography. A Blue/Orange Loading Dye, 6X, was provided.



**Figure 2.1:** 100bp DNA ladder (Promega)

### **2.10 Purification of PCR product.**

The PCR products of interest needed to be further purified in order for them to be suitable for sequencing reactions. This was achieved using Wizard® SV Gel and PCR clean-up system (Promega). This method is based on the ability of DNA to bind to silica membranes in the

presence of chaotropic salts. After electrophoresis to separate the DNA fragments, an aliquot of the PCR is added to the Membrane Binding Solution and directly purified. The system allows a choice of methods for isolation of DNA from the dissolved agarose gel slice or PCR amplification. DNA can be isolated using microcentrifugation (Eppendorf centrifuge 5415 R) to force the dissolved gel slice or PCR product through the membrane while simultaneously binding the DNA on the surface of the silica. After washing the isolated DNA fragment or PCR product, the DNA is eluted in water. The procedure was performed according to the manufacturer's instructions (appendix A9).

## 2.11 Sequencing reactions

The purified PCR samples were sequenced for mutation analysis. Sequencing was performed using the ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Perkin-Elmer). Approximately 200 ng DNA and half shots were used per reaction, (Table 2.4)

**Table 2.2:** Sequencing reactions, as described in the Perkin-Elmer ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit manual (1998).

Reagent	Quantity
Terminator Ready Reaction mix (dye terminators, dUTP <sup>*</sup> , dCTP, dATP, dITP <sup>**</sup> , AmpliTaq DNA polymerase, <i>rTh</i> pyrophosphatase, magnesium chloride, buffer)	8 $\mu$ l
DNA template	100 - 200 ng
-21 M13 Primer (forward)	1.6 pmol
dddH <sub>2</sub> O	x
Total volume	10 $\mu$ l

\* dUTP is used in place of dTTP as it results in a better T patterns because dUTP improves incorporation of T terminators.

\*\* dITP is used in place of dGTP to minimise band compressions

The reagents were vortexed then spun briefly before the PCR sequencing reaction was started (Table 2.5).

**Table 2.3:** Sequencing reaction using a Hybaid PCR Sprint Thermal Cycler

	Temperature	Time
<b>Denaturing</b>	96°C	10 sec
<b>Annealing</b>	50°C	5 sec
<b>Extension</b>	60°C	4 min

Repeat this sequence for 25 cycles, then store at 4°C before purification

## 2.12 DNA sequence analysis

Results were analysed using bioinformatic tools: The CLC DNA workbench 6 .1.1 (Inqaba biotechnology) was used to identify the sequence variants within the selected colony. ExPASy translate tool was for protein translation of the selected “candidate genes” and ClustaW2 for alignment of DNA and protein sequences to elucidate their relatedness as well as their evolutionary origin.



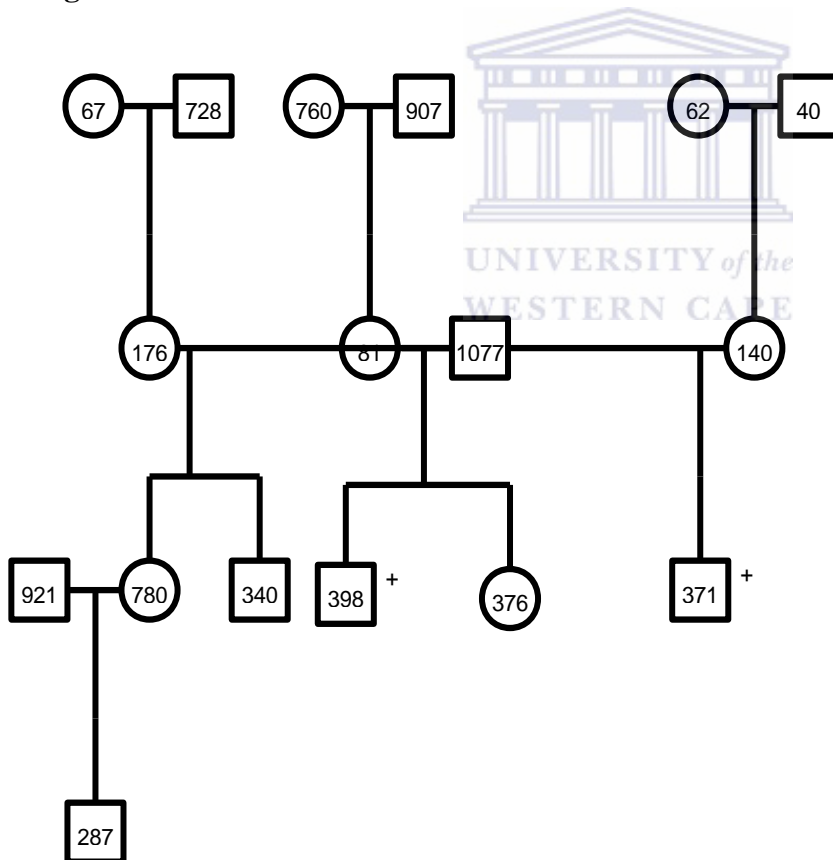
## CHAPTER THREE:

### Results

#### 3.1 Family trees of selected Vervet monkeys

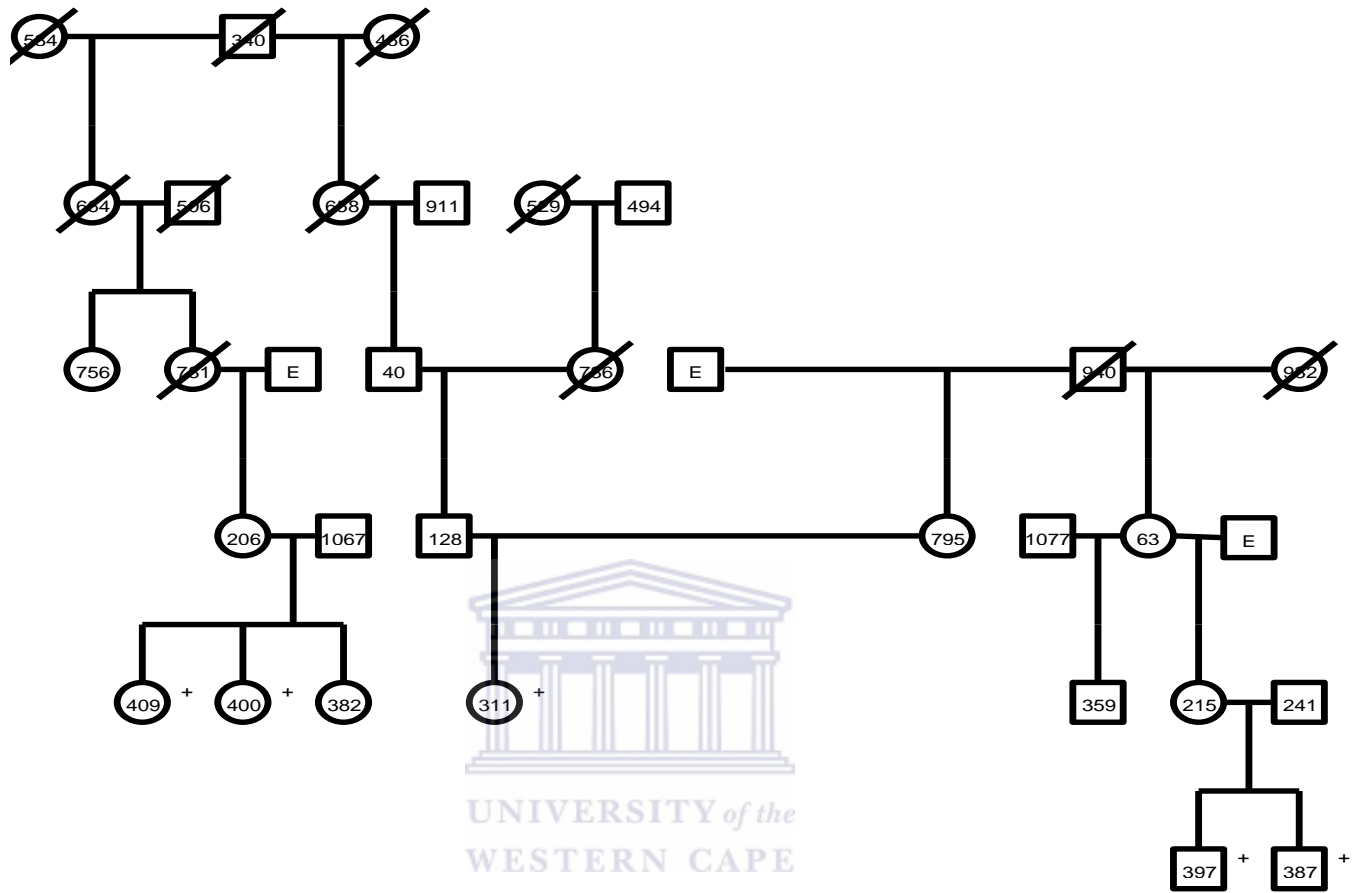
The subjects chosen for this study were members of six families, the pedigree of which are provided in Figure 3.1-3.7. Two of these contained individuals that presented with Y-sutural cataract (Figure 3.2 and 3.3).

##### Pedigree A



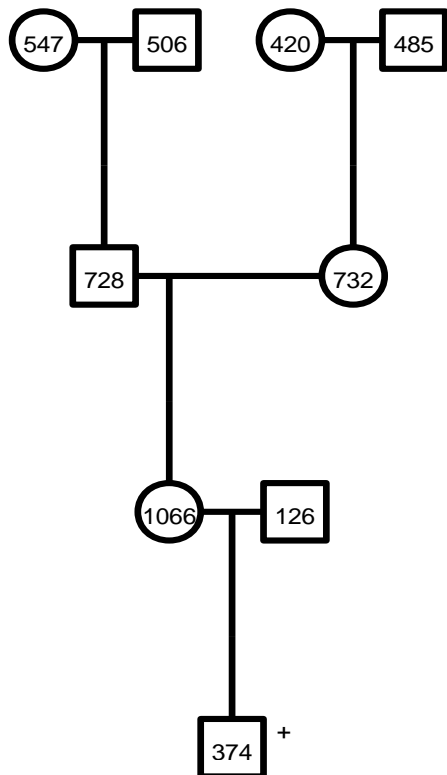
**Figure 3.1:** Pedigree A with two half-brothers with total cataract (371 and 398). The squares represent males and circles females. The cataract individuals are marked (+) respectively.

**Pedigree B**



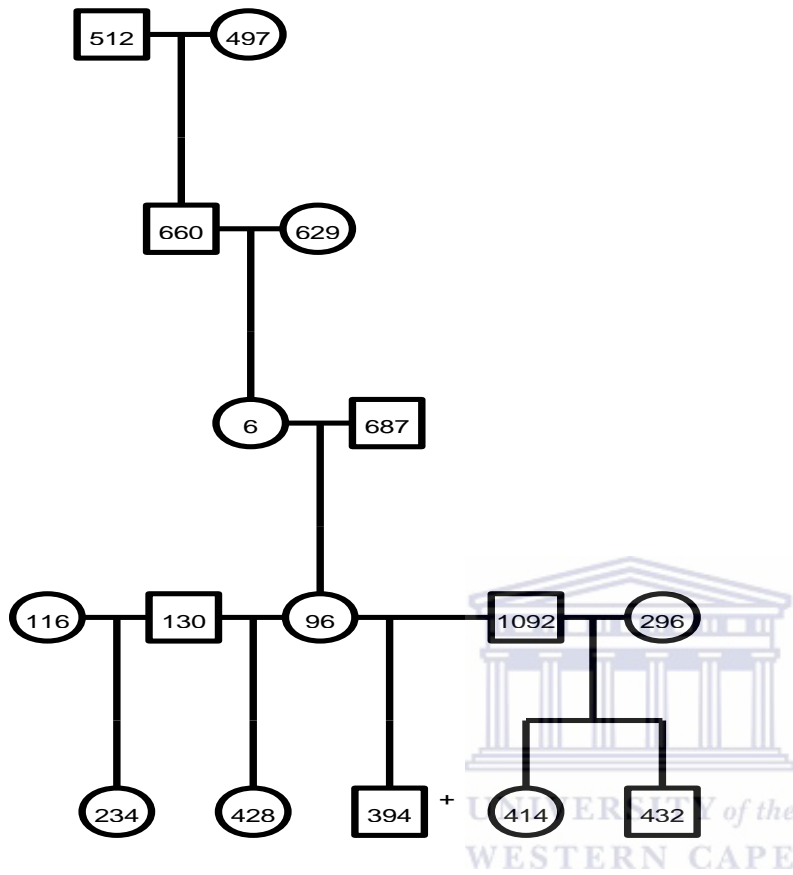
**Figure 3.2:** Pedigree B with four total cataract individuals (400, 409, 387, and 397) and one (311) Y-sutural cataract. The squares represent males, circles females, deceased with strike through (/) and the symbol E for individuals that were pair-bred in the E-cage where the parent is unknown. The cataract individuals are marked (+) respectively.

### Pedigree C



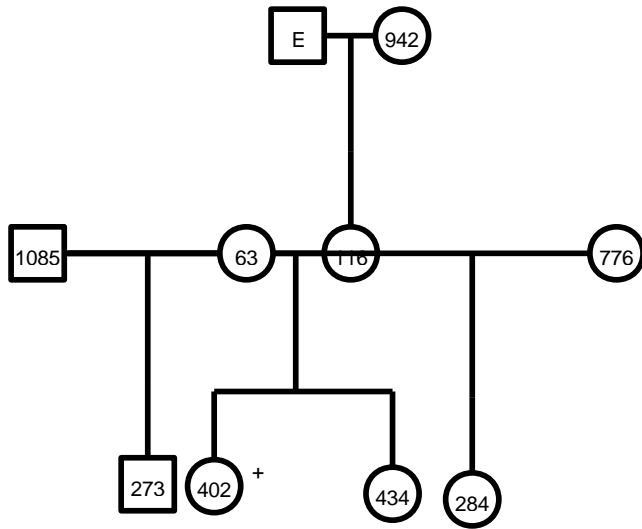
**Figure 3.3:** Pedigree C with one Y-sutural cataract individual (374). The squares represent males and circles females. The cataract individual is marked (+) respectively.

### Pedigree D

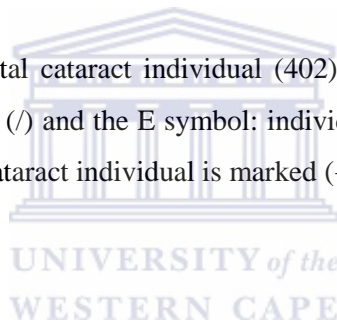


**Figure 3.4:** Pedigree D with one total cataract individual (394). The squares represent males and circles females. The cataract individual is marked (+) respectively.

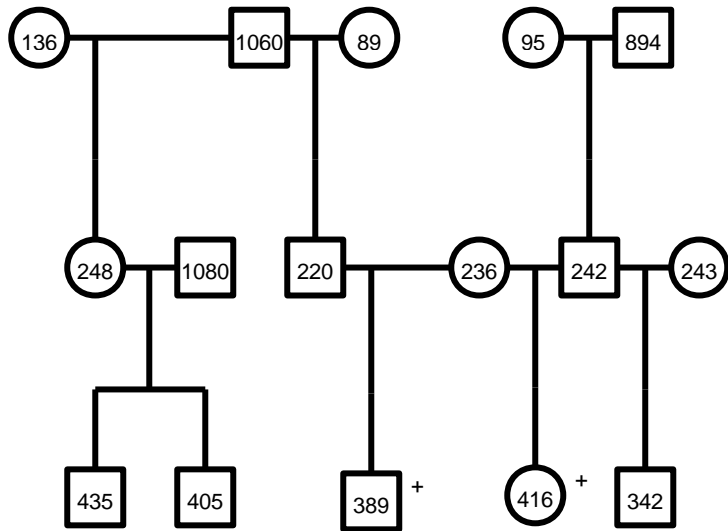
### Pedigree E



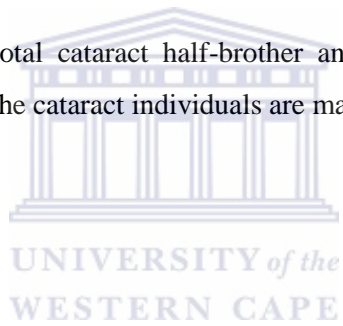
**Figure 3.5:** Pedigree E with one total cataract individual (402). The squares represent males, circles females, deceased with strike through (/) and the E symbol: individuals that were pair-bred in the E-cage where the father was unknown. The cataract individual is marked (+) respectively.



**Pedigree F**



**Figure 3.6:** Pedigree F with two total cataract half-brother and sister (389 and 416). The squares represent males and circles females. The cataract individuals are marked (+) respectively.



**A**



**B**



**Figure 3.7:** A) One of the two monkeys with Y-sutural cataract. B) One of the ten monkeys with total cataract. The above morphologies were unilateral or bilateral in some individuals.

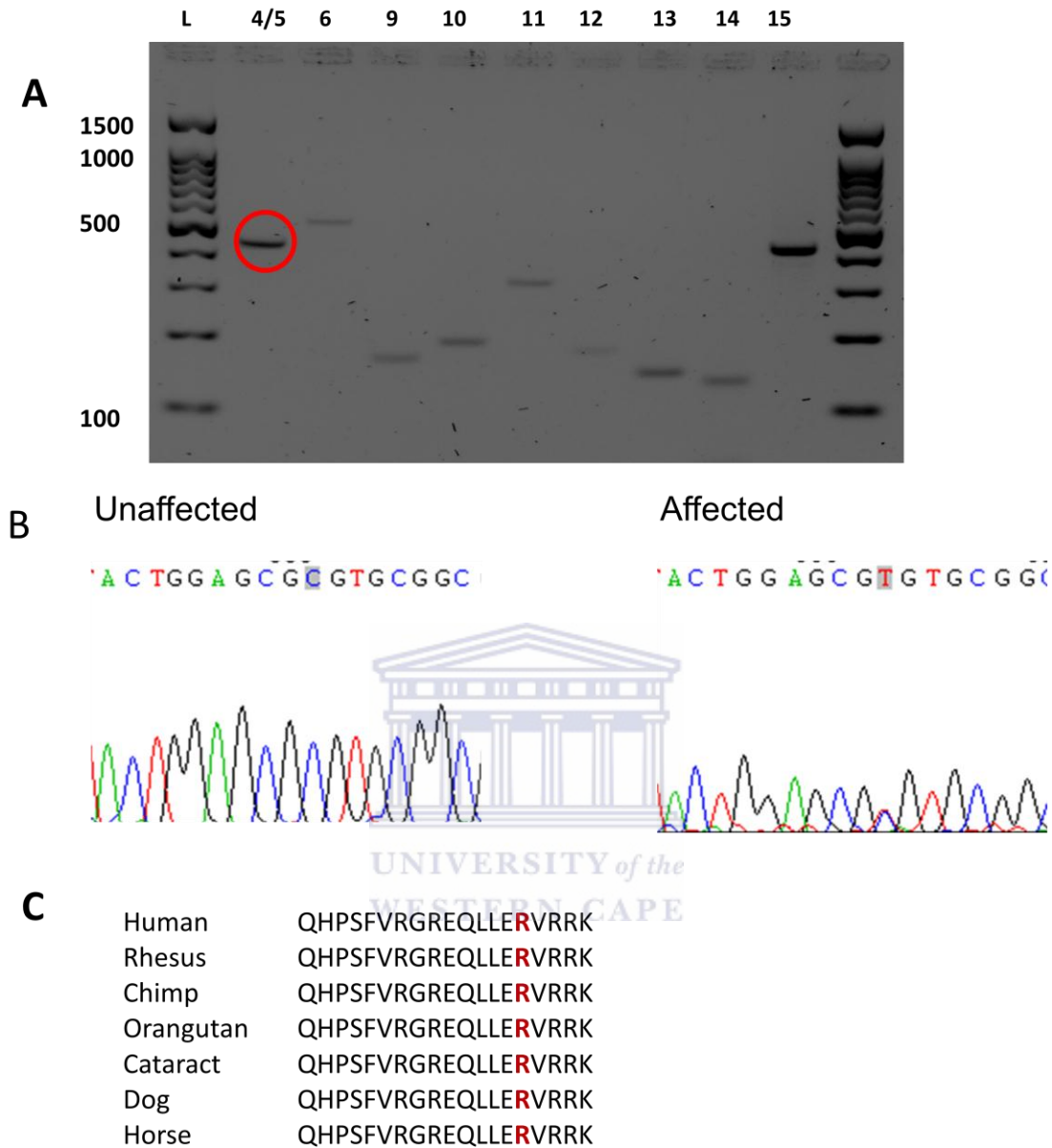
### **3.2 DNA sequence variations in captive bred vervet monkeys.**

All the selected genes were screened, sequenced and analysed. The optimized PCR conditions for all selected primers are shown in Appendix B. Reported below are only the coding exons that revealed mutations in the prioritized genes.

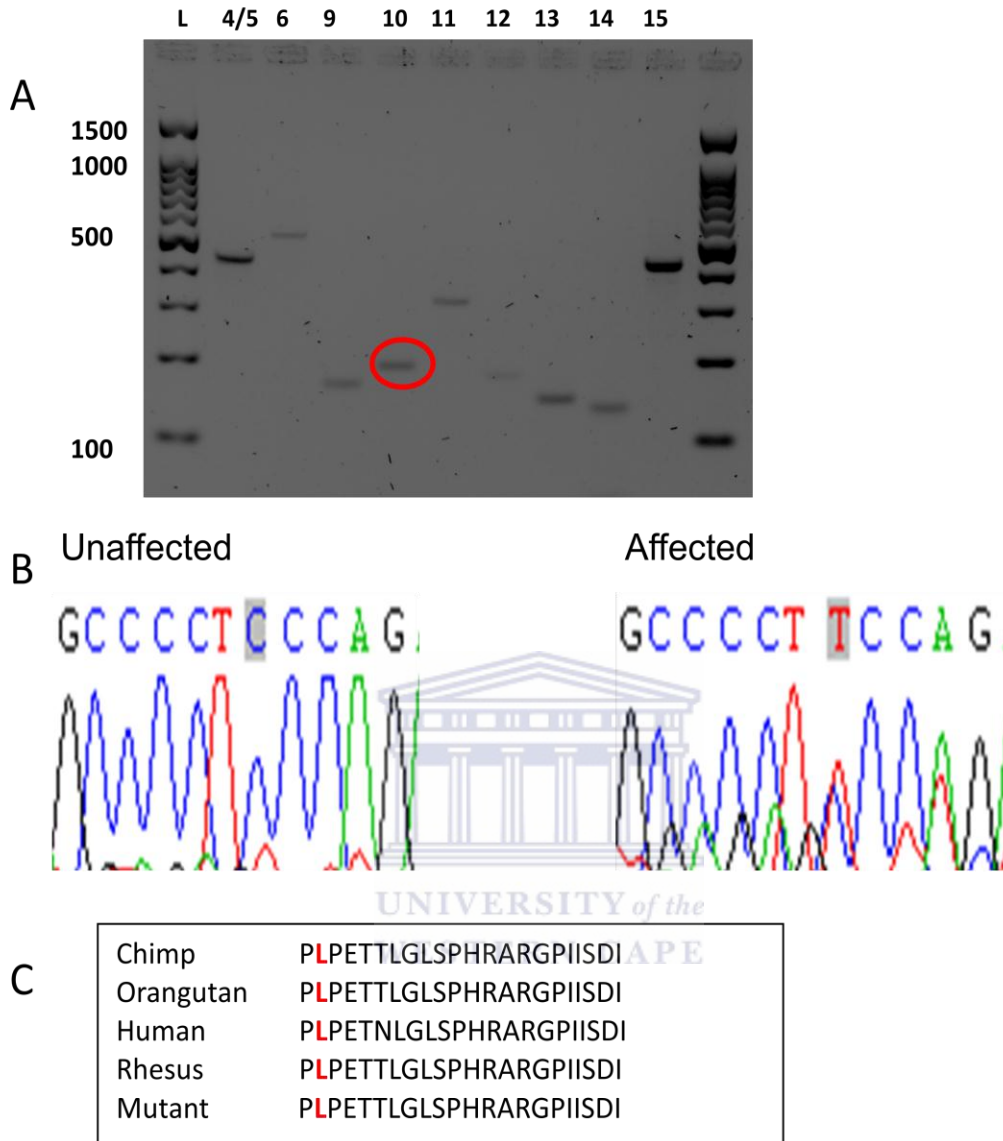
#### **3.2.1 DNA Sequence variants in *HSF4***

To detect *HSF4* sequence abnormalities, all the coding exons were screened by direct genomic PCR sequencing. Controls and carriers were sequenced in order to identify common variants in *HSF4*. Three sequence variants were identified which were not present in control individuals. The findings include two transition silent mutations: c.1313 C>T found in eight cataract and three carriers (R116>R) (Figure 3.8) in exon 5, c. 4095C>T in one cataract individual (L245>L) (Figure 3.9) in exon 10. Four cataract and two carriers were affected with c.5900C>T which resulted in amino acid change P421>L (Figure 3.10) in exon 14. The regions in which these mutations occurred within *HSF4* are illustrated in Figure 3.11 respectively.

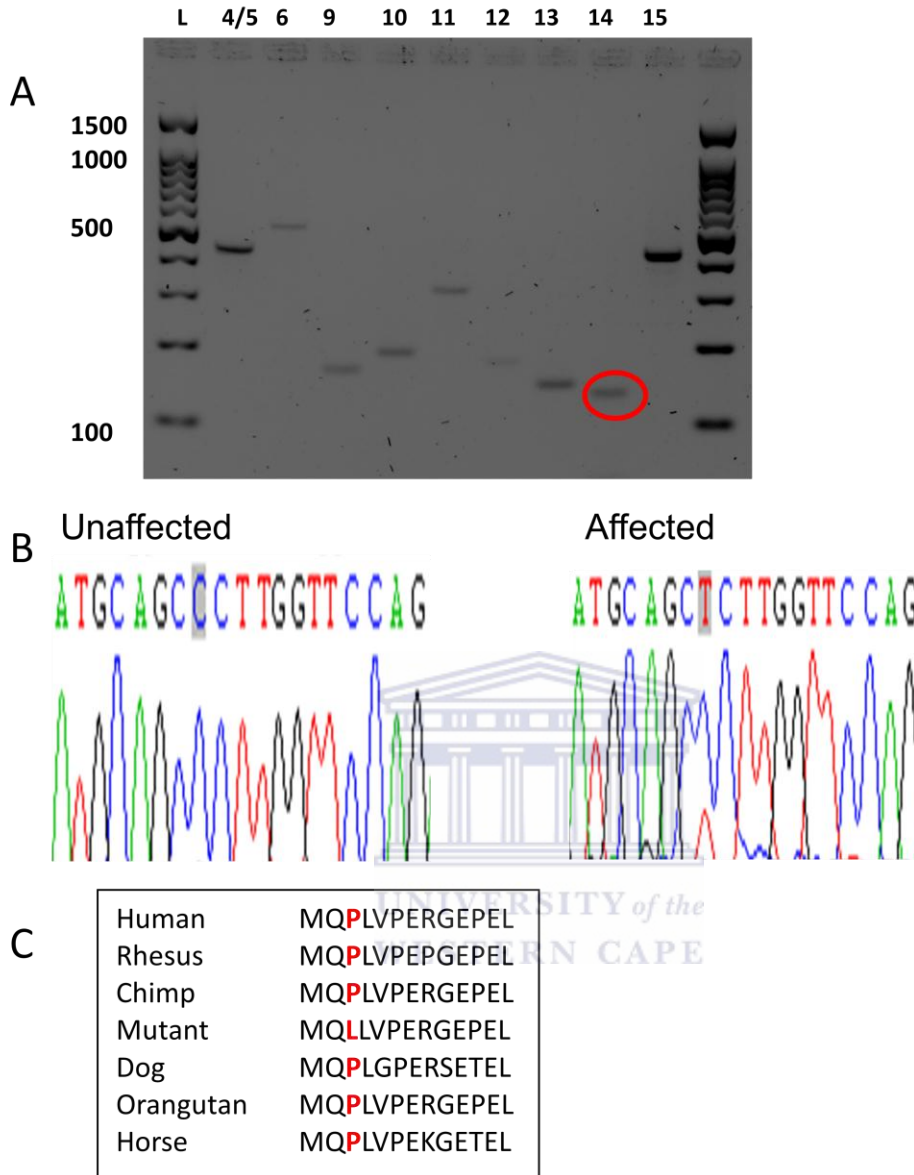




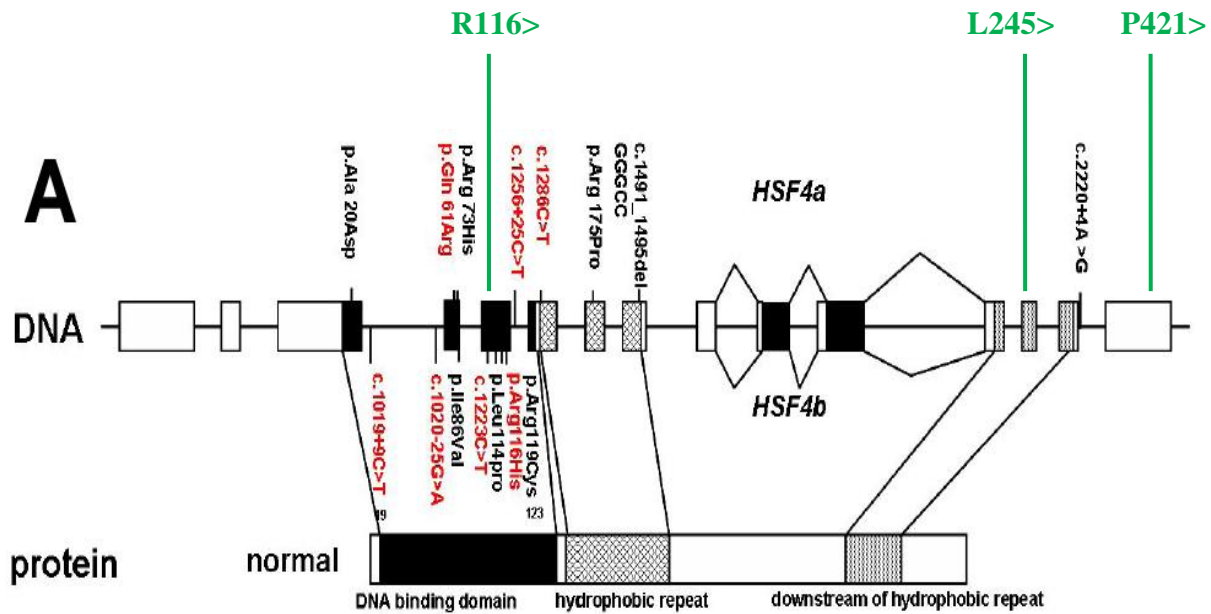
**Figure 3.8:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of HSF4 gene with one mutation in exon 5 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1313 C>T) found in exon 5 *HSF4*. The change at codon 116 (R116>R) was found in eight cataract and three carriers. **C)** The protein sequence alignment among different species where cataract represents for Vervet.



**Figure 3.9:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *HSF4* gene with one mutation in exon 10 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1700 C>T) found in exon 10 *HSF4*. The change at codon 245 (L245L) was found in one affected individual. **C)** The protein sequence alignment among different species where mutant represents Vervet.



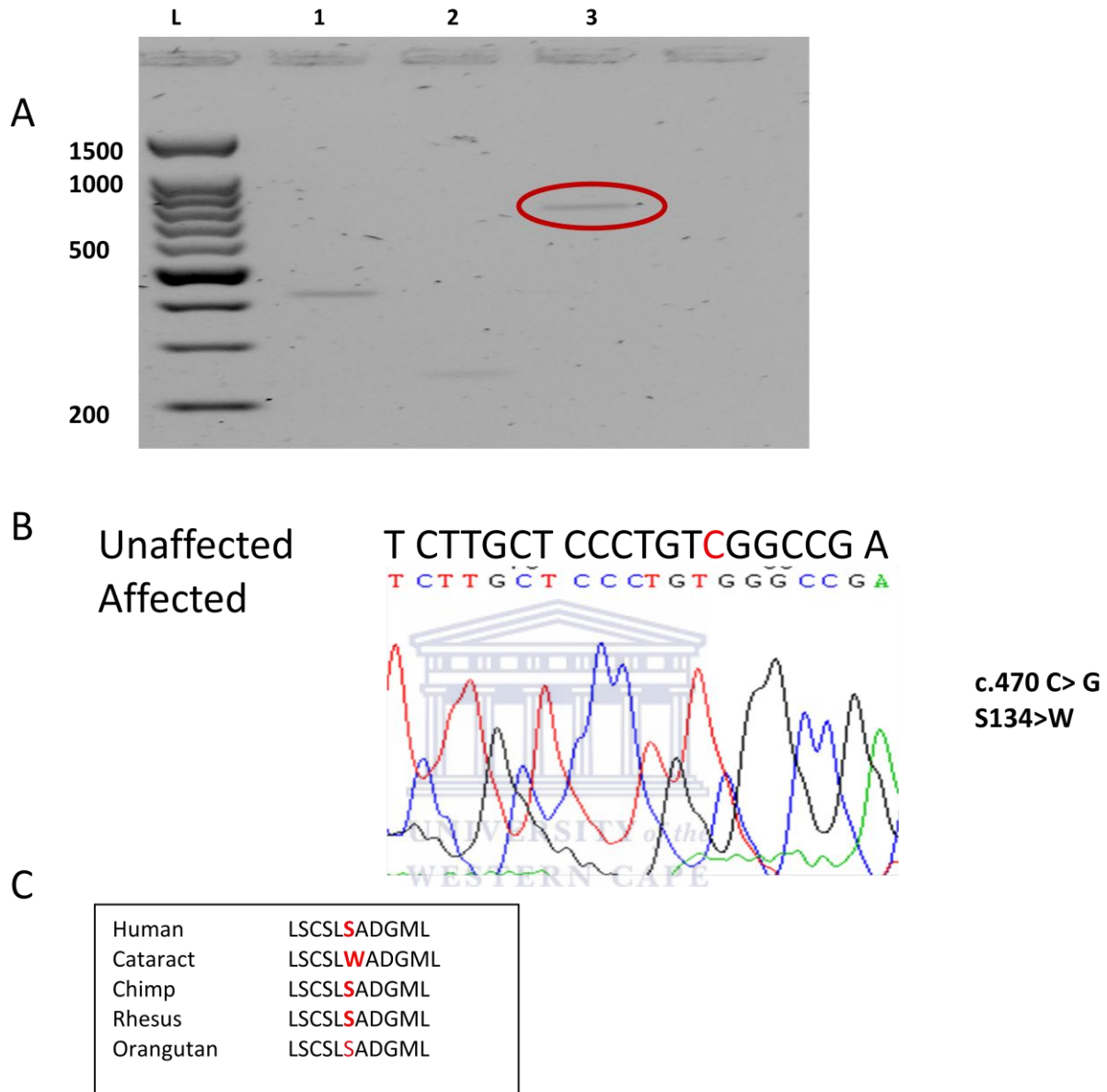
**Figure 3.10:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *HSF4* gene with sequence variant in exon 14 (red circle). **B)** Sequence chromatogram showing transition missense mutation (c.5900C>T) found in exon 14 *HSF4*. The change at codon 421 was found in four cataract and two carriers, this resulted in the substitution of Proline to Leucine (P421L). **C)** The protein sequence alignment among different species where mutant represents Vervet.



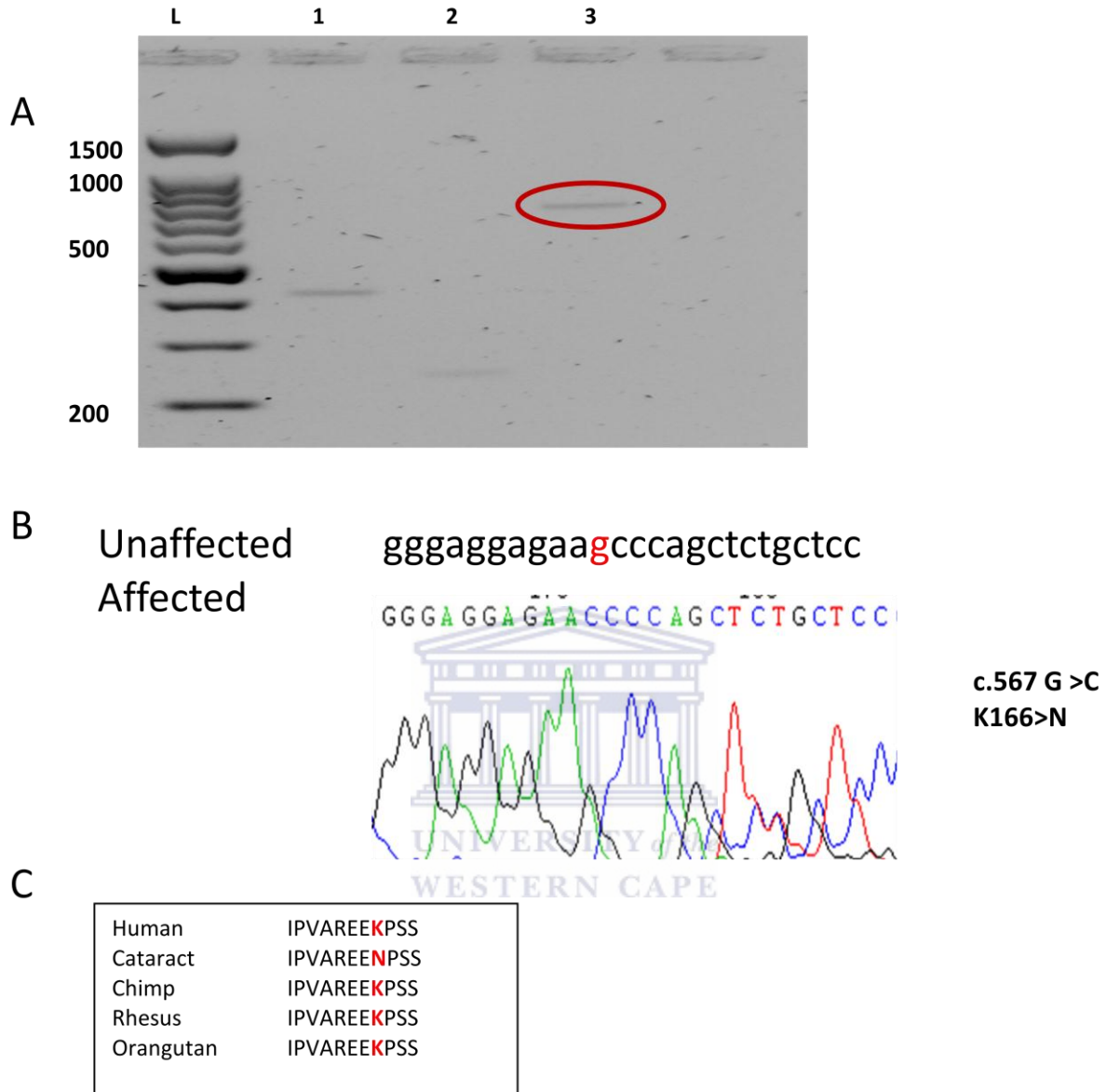
**Figure 3.11:** The structure of *HSF4* with previously reported mutations. Adapted from Shi *et al.*, 2008. The mutations that were identified in Vervet monkeys are indicated with green.

### 3.2.2 DNA sequence variants in *CRYAA*

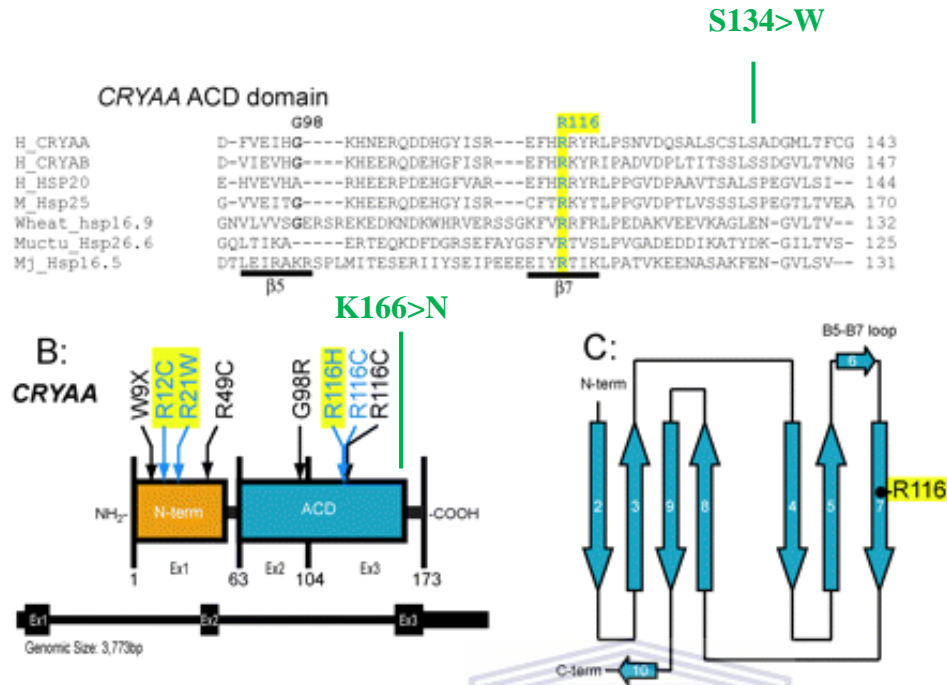
To detect for sequence abnormalities in *CRYAA* gene, all the coding exons were screened following the same procedure as *HSF4*. Sequencing of exon 3 showed two transversion missense mutations: c.470C>G was observed in two Y-sutural, one carrier and three total cataract individuals. This missense mutation resulted in an amino acid change S134>W (Figure 3.12). The second mutation, c. 567G>C resulted in amino acid change K166>N (Figure 3.13) and was observed in one of the Y-sutural and two total cataract monkeys. These sequence variants were not observed in the controls. The regions of *CRYAA* in which these mutations were identified are shown in figure 3.14 respectively.



**Figure 3.12:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *CRYAA* gene with one mutation in exon 3 (red circle). **B)** Sequence chromatogram showing transversion missense mutation (c.470 C > G) found in exon 3 of *CRYAA*. The change at codon 134 was found in one carrier and five affected individuals, this resulted in substitution of Serine to Tryptophan (S134>W). **C)** The protein sequence alignment among different species where cataract represents Vervet.



**Figure 3.13:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *CRYAA* gene with one mutation in exon 3 (red circle). **B)** Sequence chromatogram showing transversion missense mutation (c.567 G > C) found in exon 3 of *CRYAA*. The change at codon 166 was found in three affected individuals, this resulted in substitution of Lysine to Asparagine (K166>N). **C)** The protein sequence alignment among different species where cataract represent Vervet.



**Figure 3.14:** The structure of alpha crystalline genes with previously reported mutations. Adapted from Hansen *et al.*, 2007. The mutations that were identified in Vervet monkeys are indicated in green.

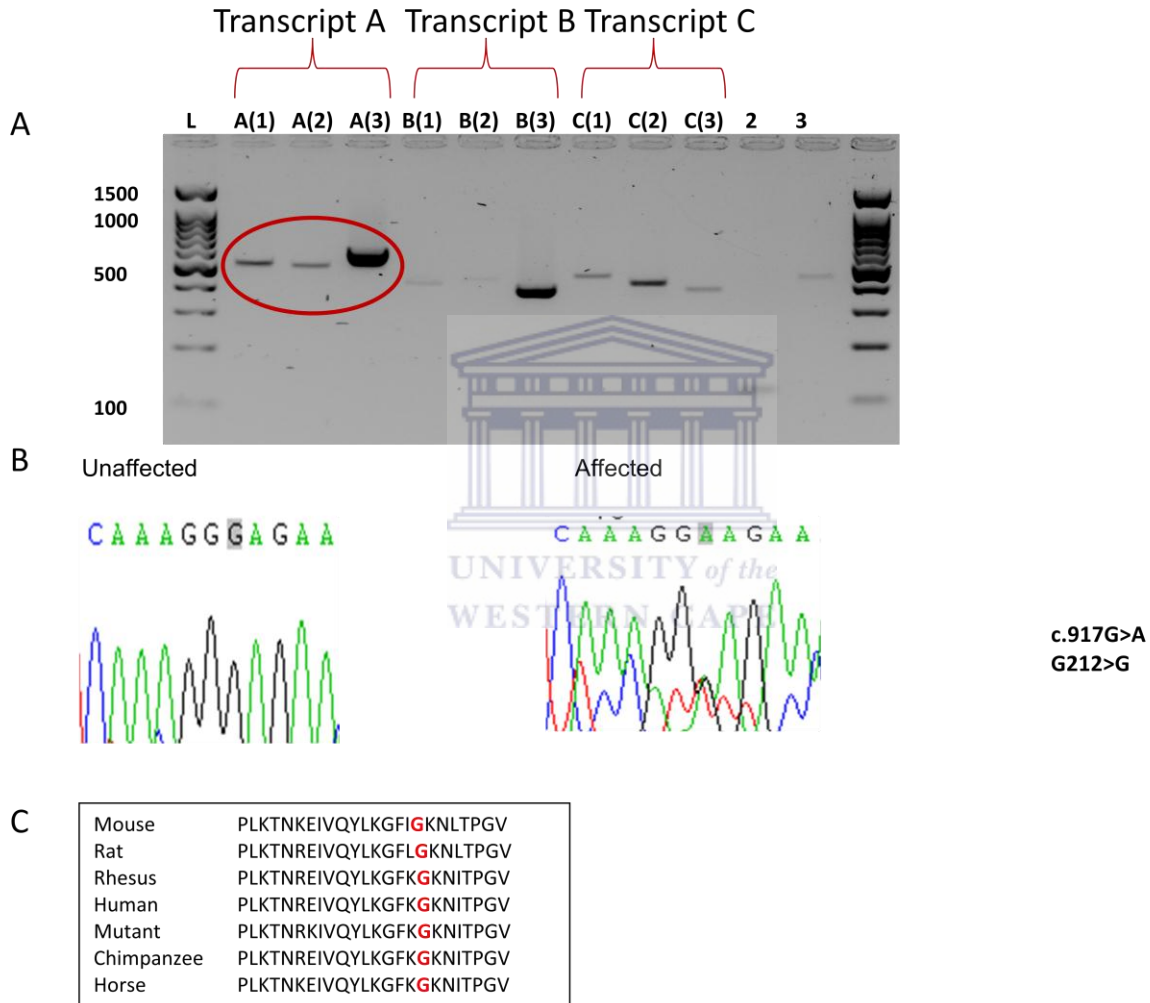
UNIVERSITY of the  
WESTERN CAPE

### 3.2.3 DNA sequence variants in *GCNT2*

The *GCNT2* transcripts showed twelve sequence variants. Transcript A showed three silent and one missense mutation: Four carriers and eight cataract individuals were sharing G212G (Figure 3.15), one carrier and three cataract individuals with H256>H (Figure 3.16), one carrier and three cataract individuals had M258>V (Figure 3.17), three carriers and six cataracts subjects had N275>N (Figure 3.18). Two missense mutations were identified in transcript B: one carrier and four individuals were affected with (V16>I) (Figure 3.19), two carriers and three cataract individuals had (Y122>F) (Figure 3.20). Transcript C showed four transition silent and one missense mutations: two cataract individuals had S15>S (Figure 3.21), one with S24>N (Figure

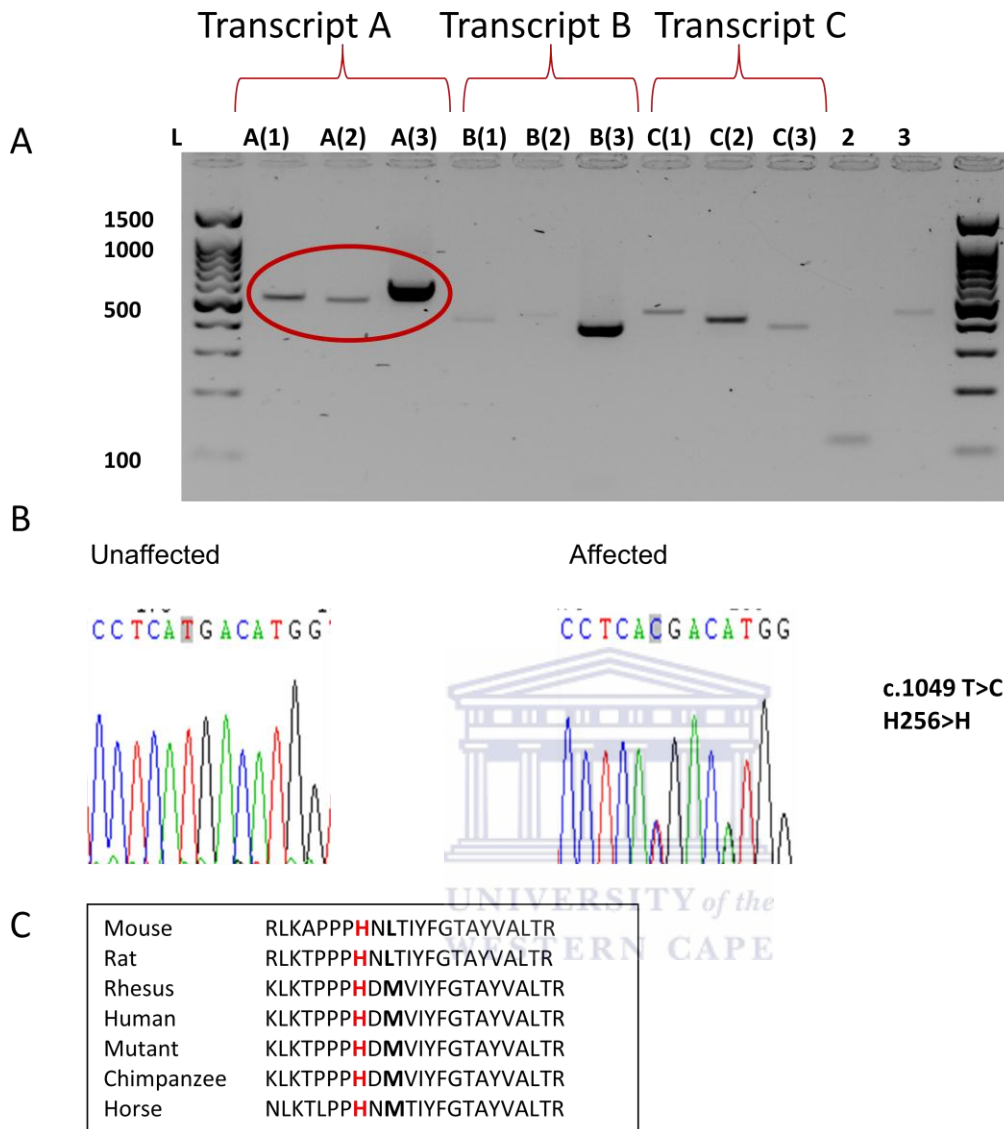


3.22) three had S38>S (Figure 3.23), five with I118I (Figure 3.24) and seven with D194>D (Figure 3.25). Exon three revealed a silent mutation Y373>Y (Figure 3.26) in seven cataract subjects. The predicted regions of the above mutations are illustrated in Figure 3.27. These mutations were not found in all the controls.

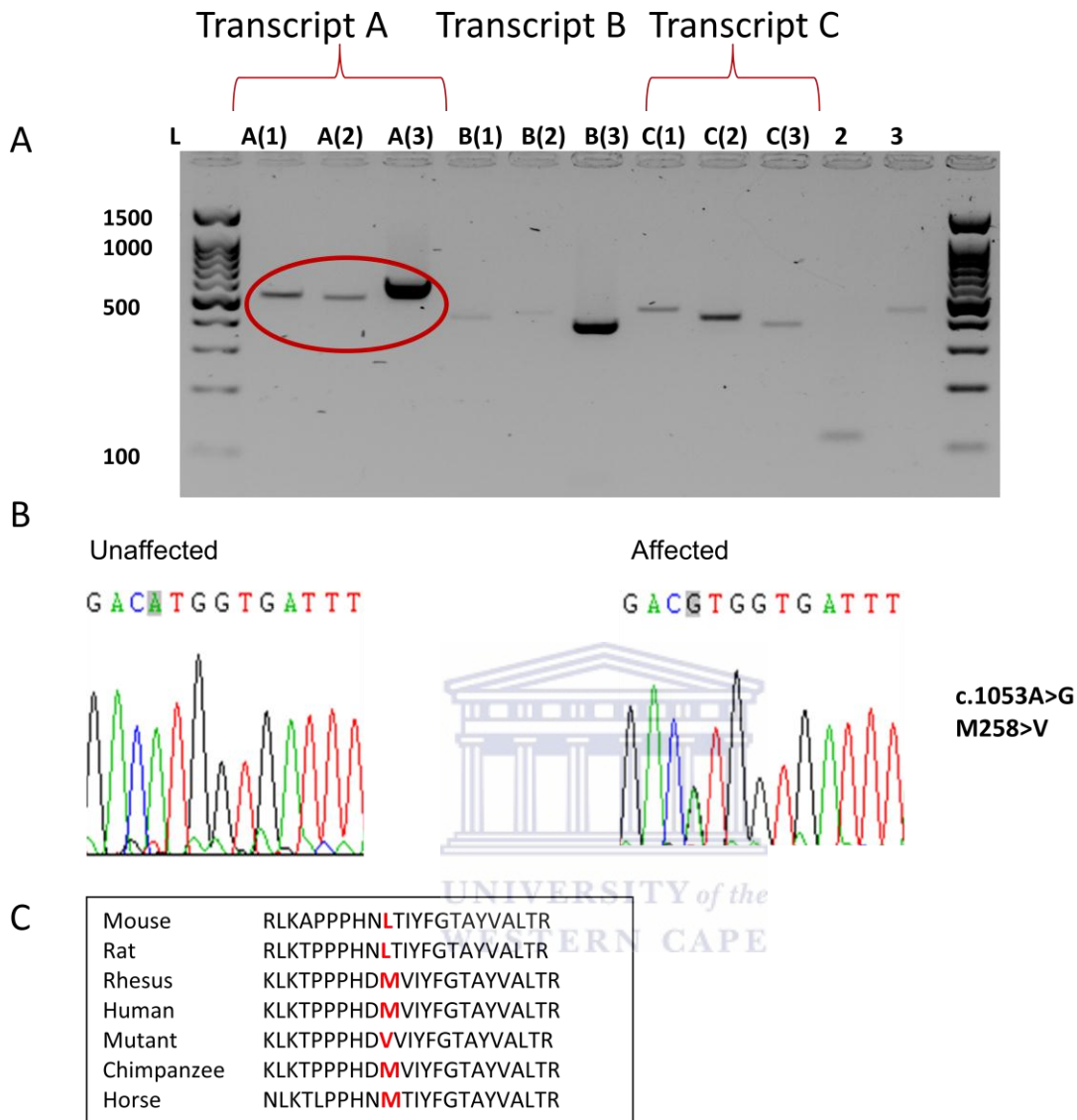


**Figure 3.15:** Genetic analysis of selected captive-bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.917G>A) found in exon 1 of *GCNT2* transcript A. The change at codon 212 (G212>G) was found in three carriers and eight affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

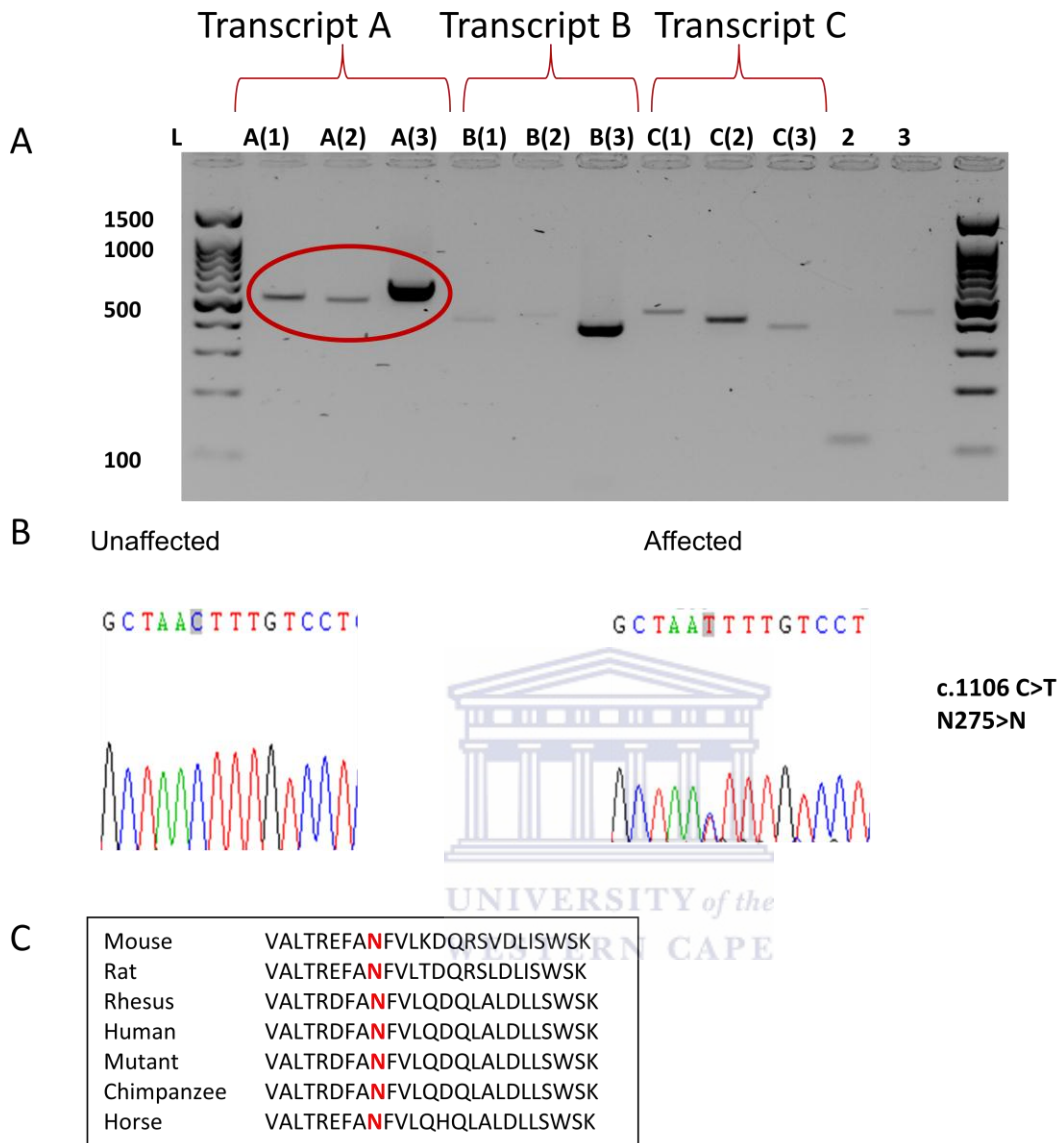




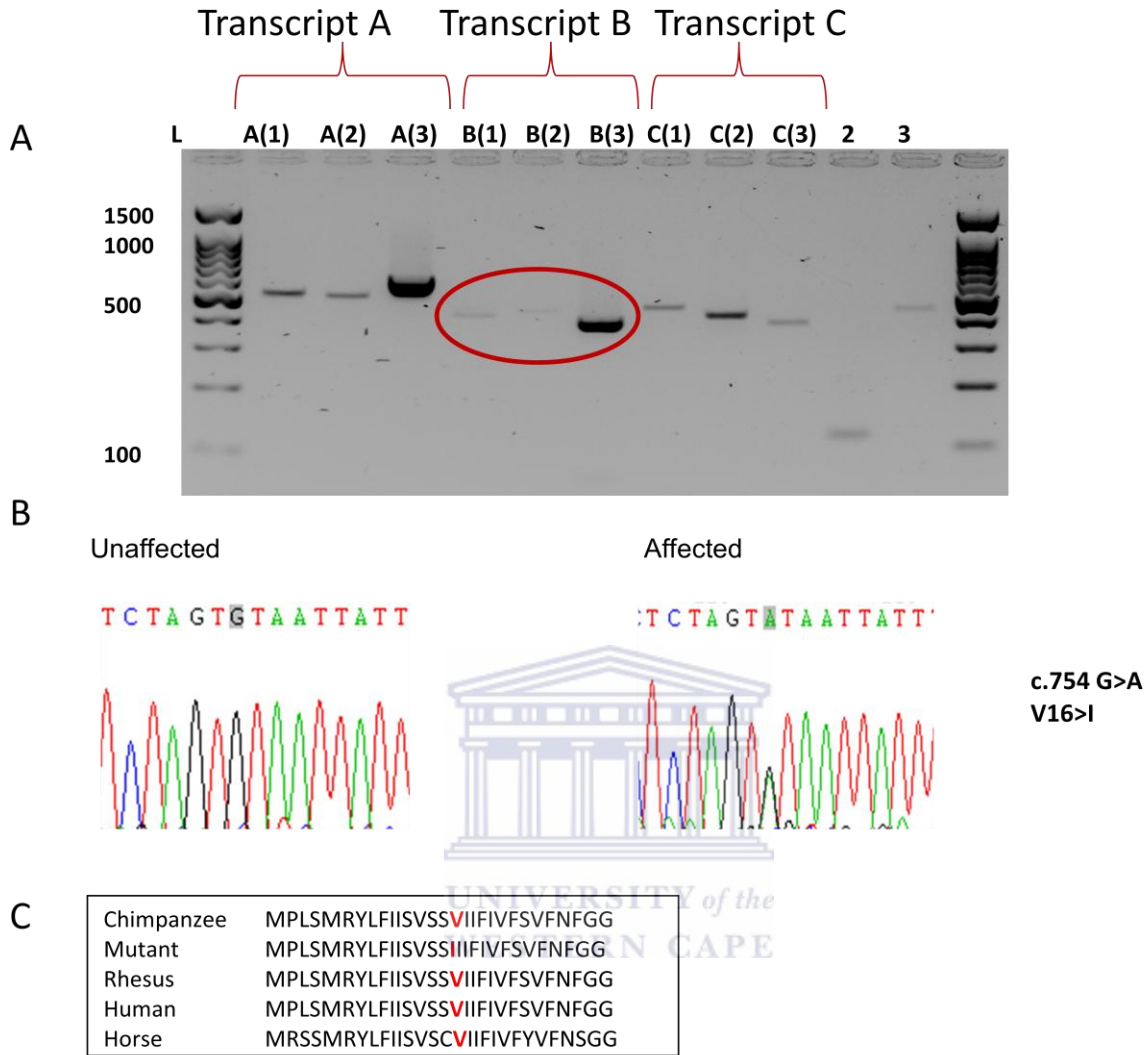
**Figure 3.16:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1 (red circle). **B)** Sequence chromatogram showing transition silent mutation (c.1049 T>C) found in exon 1 of *GCNT2* transcript A. The change at codon 256 (H256>H) was found in one carrier and three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



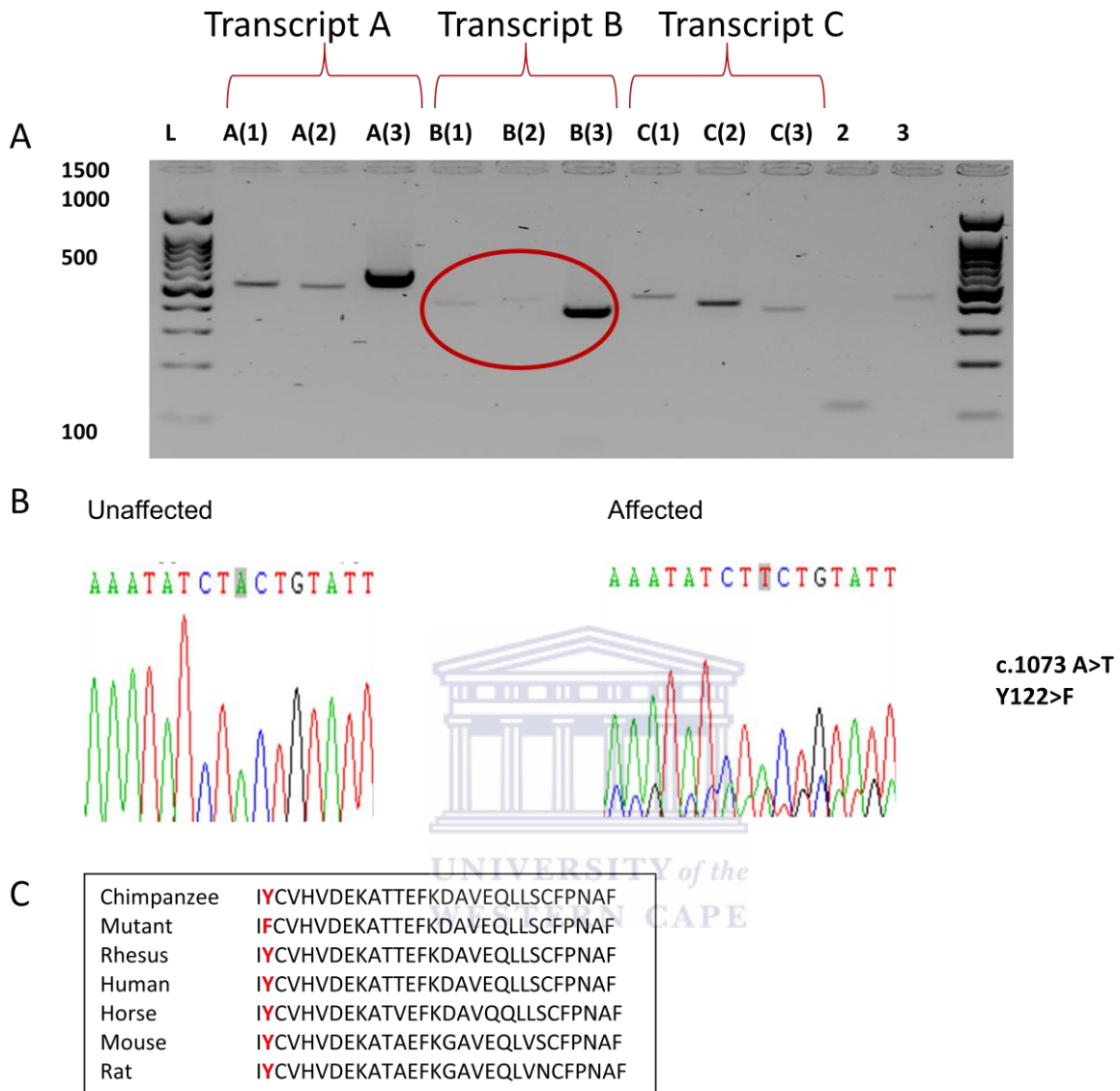
**Figure 3.17:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition missense mutation (c.1053A>G) found in exon 1 of *GCNT2* transcript A. The change at codon 258 (M258>V) was found in one carrier and three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



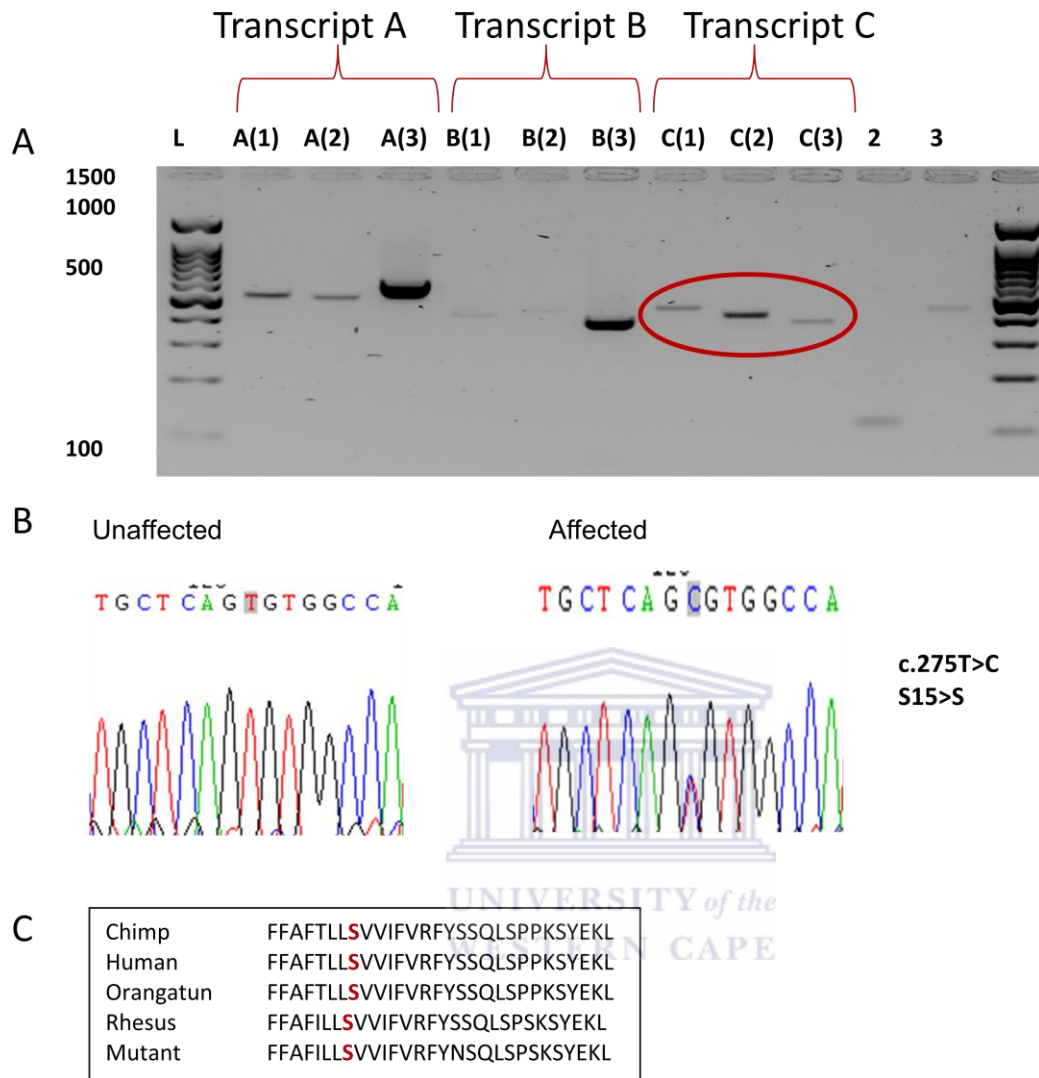
**Figure 3.18:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition silent mutation (c.1106 C>T) found in exon 1 of *GCNT2* transcript A. The change at codon 275 (N275>N) was found in three carriers and six affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



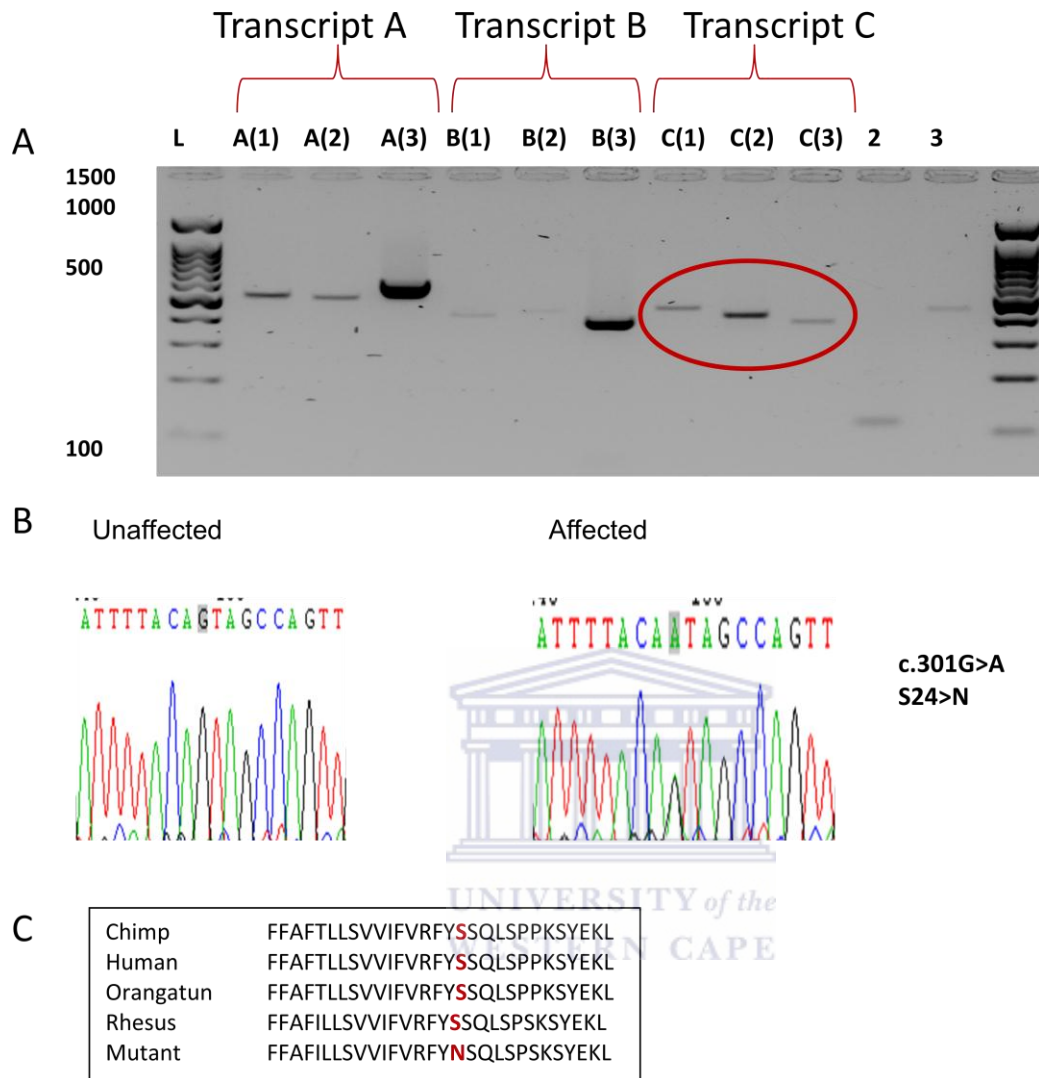
**Figure 3.19:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition missense mutation (c.754 G>A) found in exon 1 of *GCNT2* transcript B. The change at codon 16 (V16>I) was found in one carrier and four affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



**Figure 3.20:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transversion missense mutation (c.1073 A>T) found in exon 1 of *GCNT2* transcript B. The change at codon 122 (Y122>F) was found in two carriers and three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

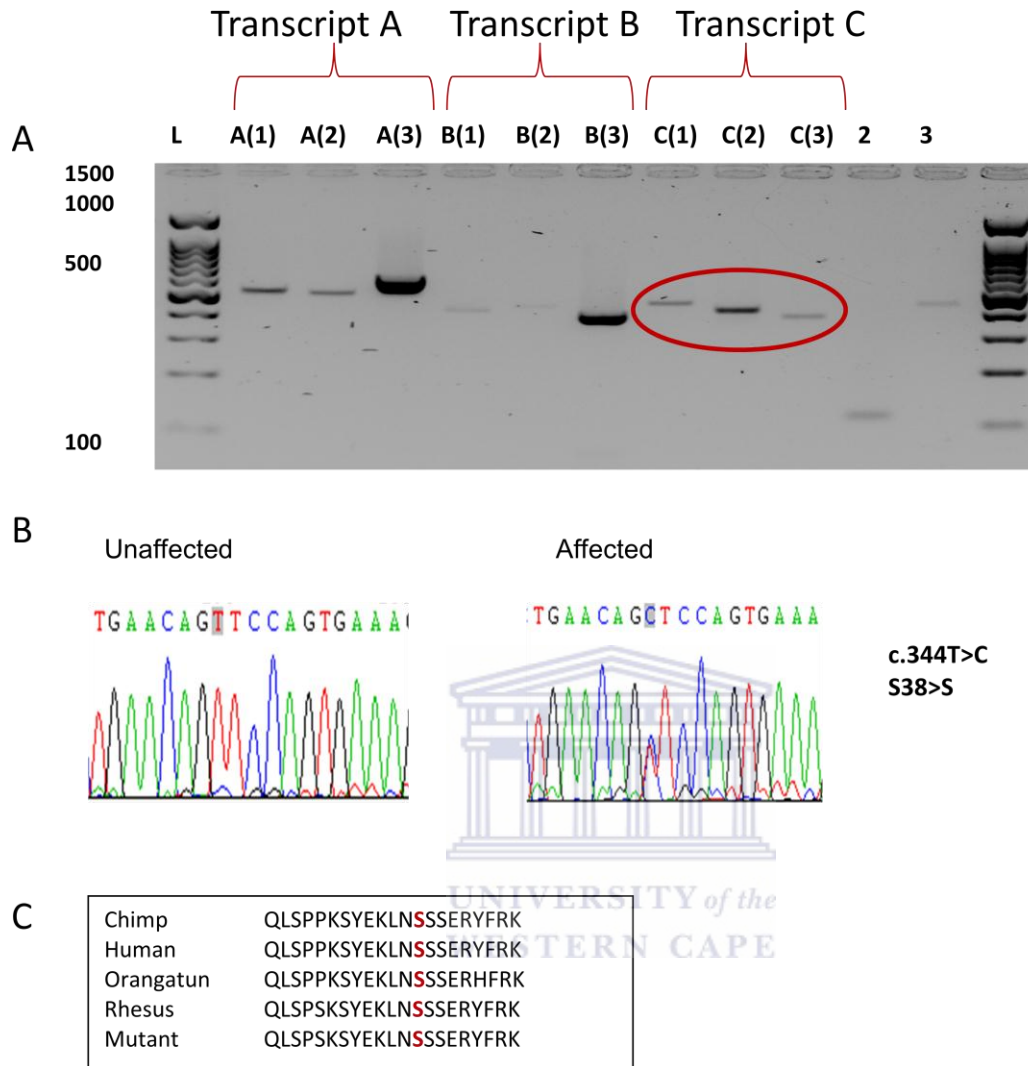


**Figure 3.21:** Genetic analysis of selected captive- bred vervet monkeys. A) A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. B) Sequence chromatogram showing transition silent mutation (c.275T>C) found in exon 1 of *GCNT2* transcript C. The change at codon 15 (S15>S) was found in two affected individuals. C) The protein sequence alignment among different species where mutant represents Vervet.



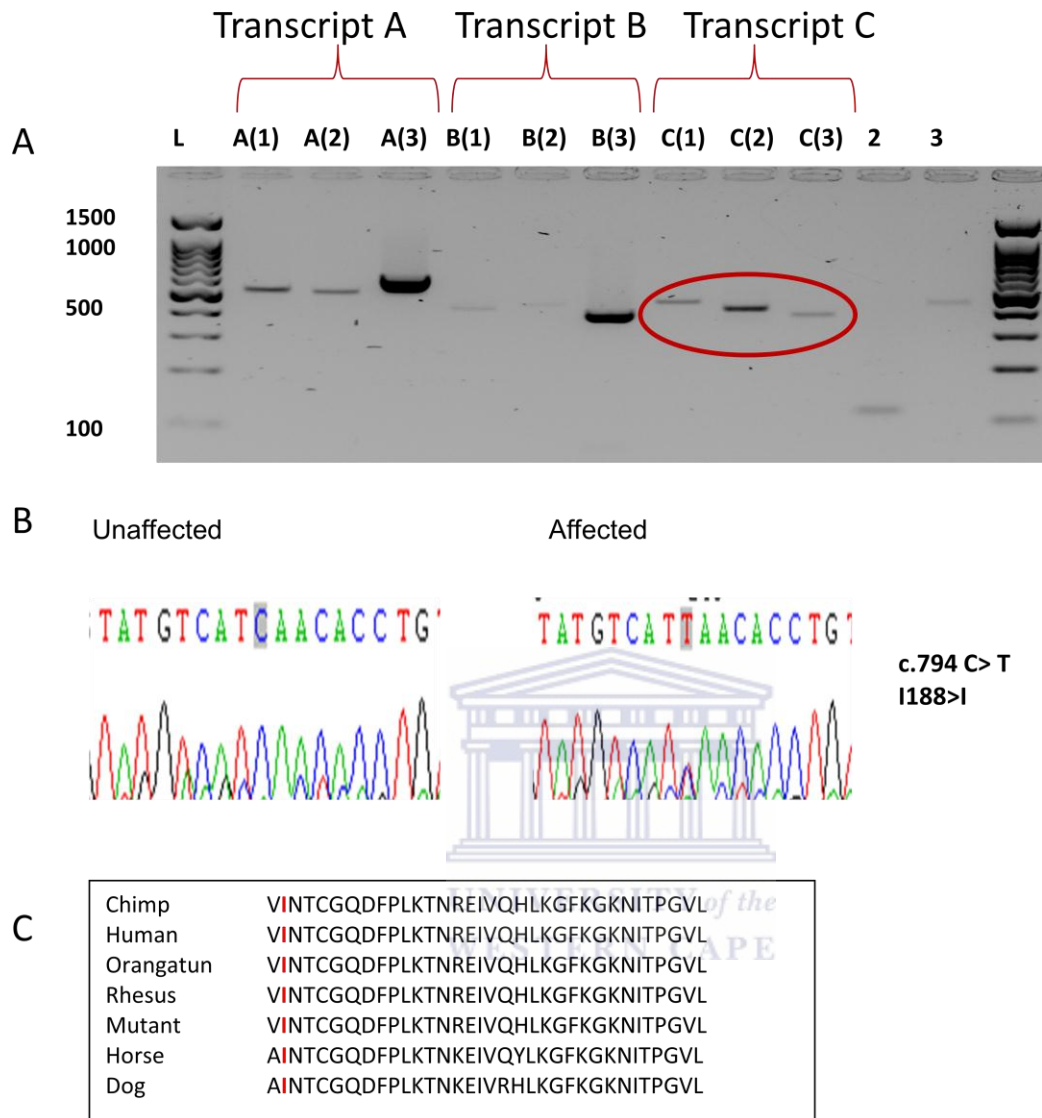
**Figure 3.22:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition missense mutation (c.301G>A) found in exon 1 of *GCNT2* transcript C. The change at codon 24 was found in one affected individual, this resulted in the substitution of Serine to Asparagine (S24>N). **C)** The protein sequence alignment among different species where mutant represents Vervet.



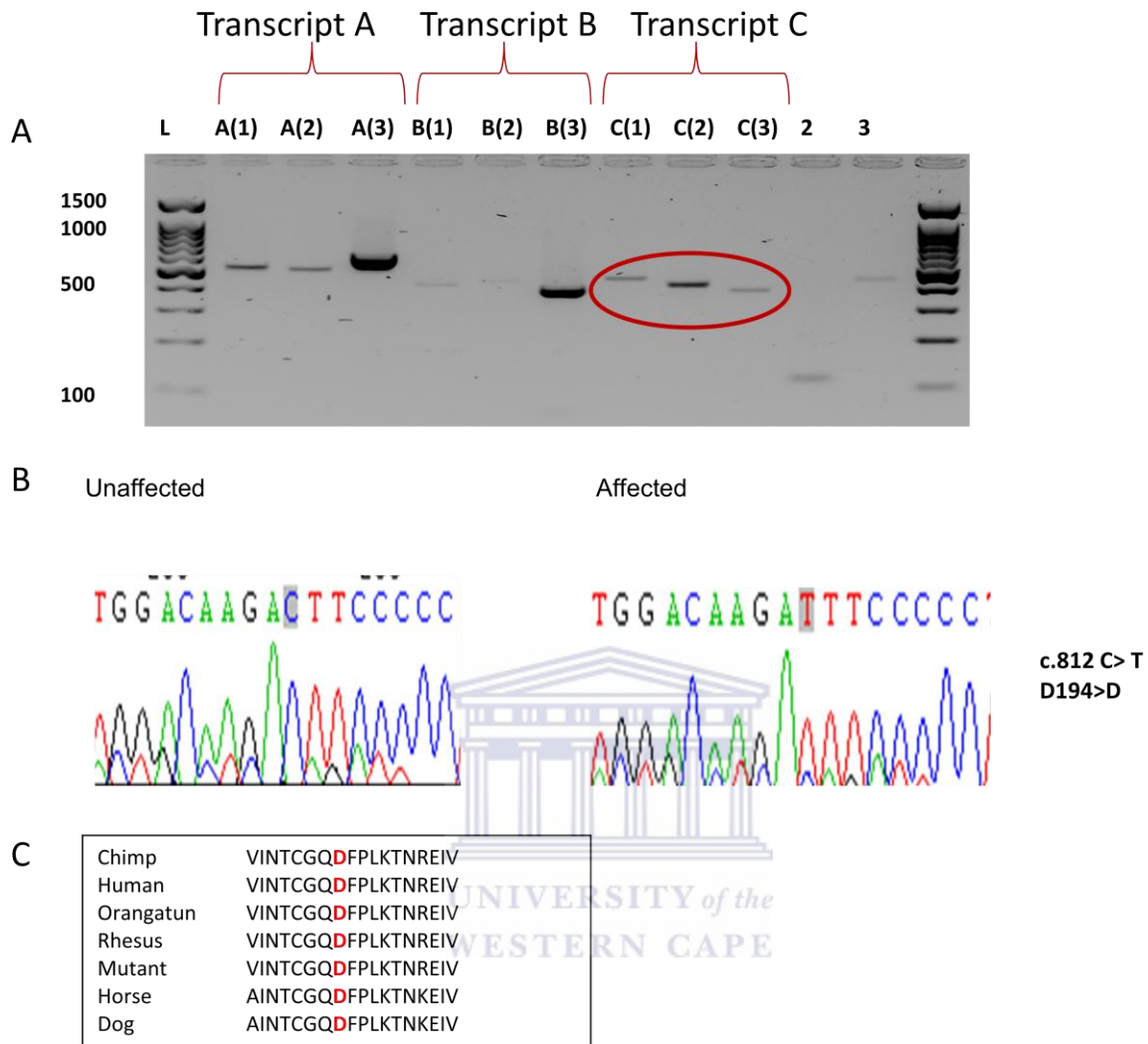


**Figure 3.23:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition silent mutation (c.344T>C) found in exon 1 of *GCNT2* transcript C. The change at codon 38 (S38>S) was found in three affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

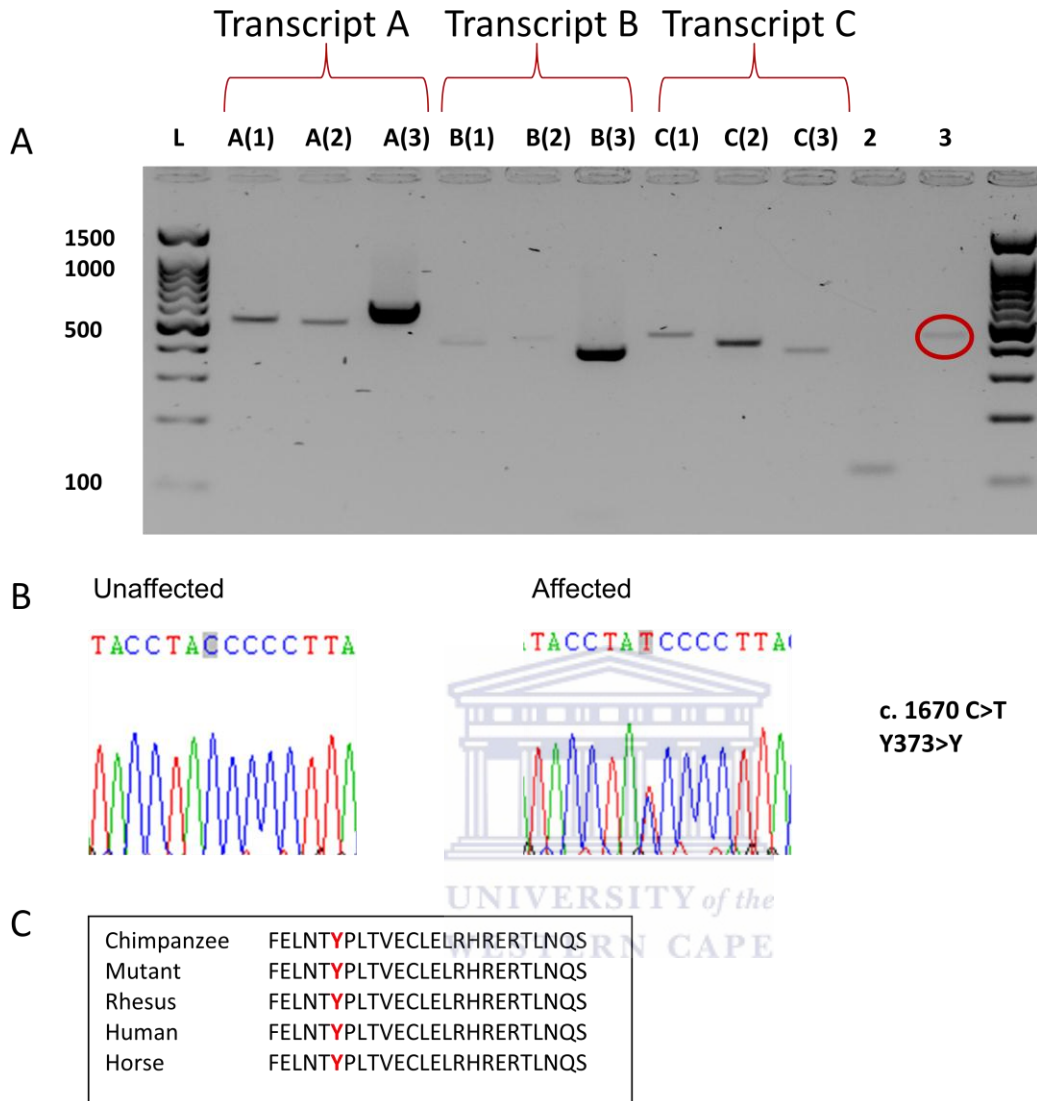




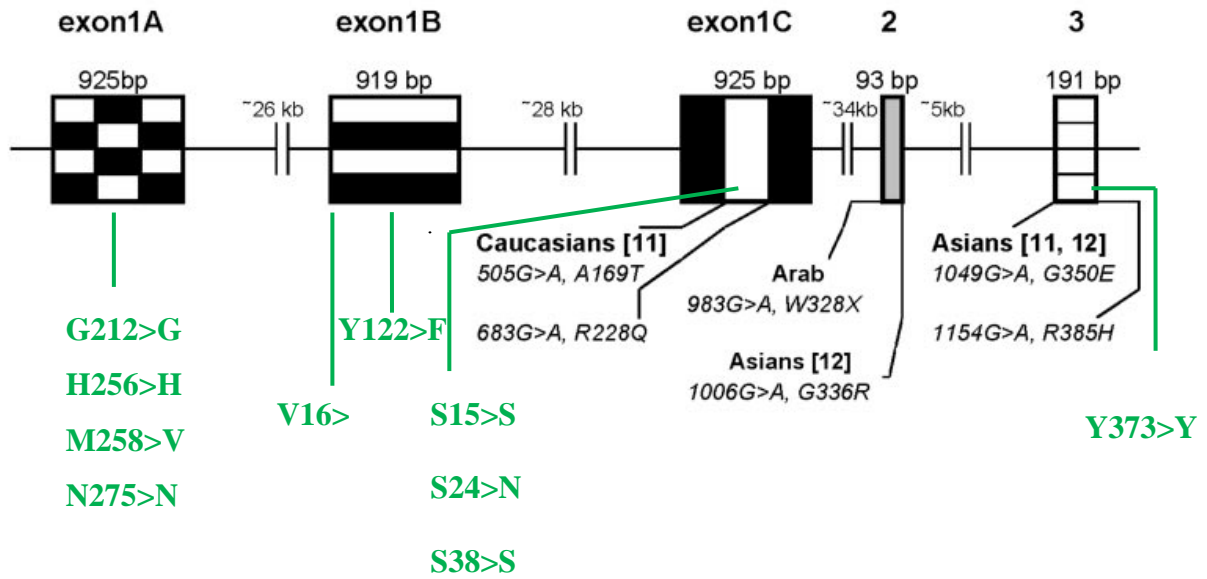
**Figure 3.24:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition silent mutation (c.794 C> T) found in exon 1 of *GCNT2* transcript C. The change at codon 188 (I188>I) was found in five affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



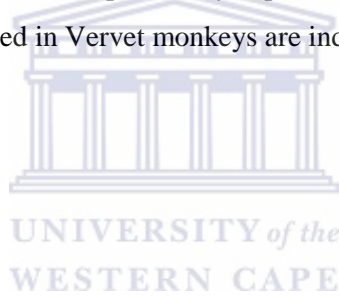
**Figure 3.25:** Genetic analysis of selected captive-bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition silent mutation (c.812 C> T) found in exon 1 of *GCNT2* transcript C. The change at codon 194 (D194>D) was found in seven affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.



**Figure 3.26:** Genetic analysis of selected captive- bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *GCNT2* gene with one mutation in exon 1. **B)** Sequence chromatogram showing transition silent mutation (c.1670C> T) found in exon 3 of *GCNT2*. The change at codon 373 (Y373>Y) was found in six affected individuals. **C)** The protein sequence alignment among different species where mutant represents Vervet.

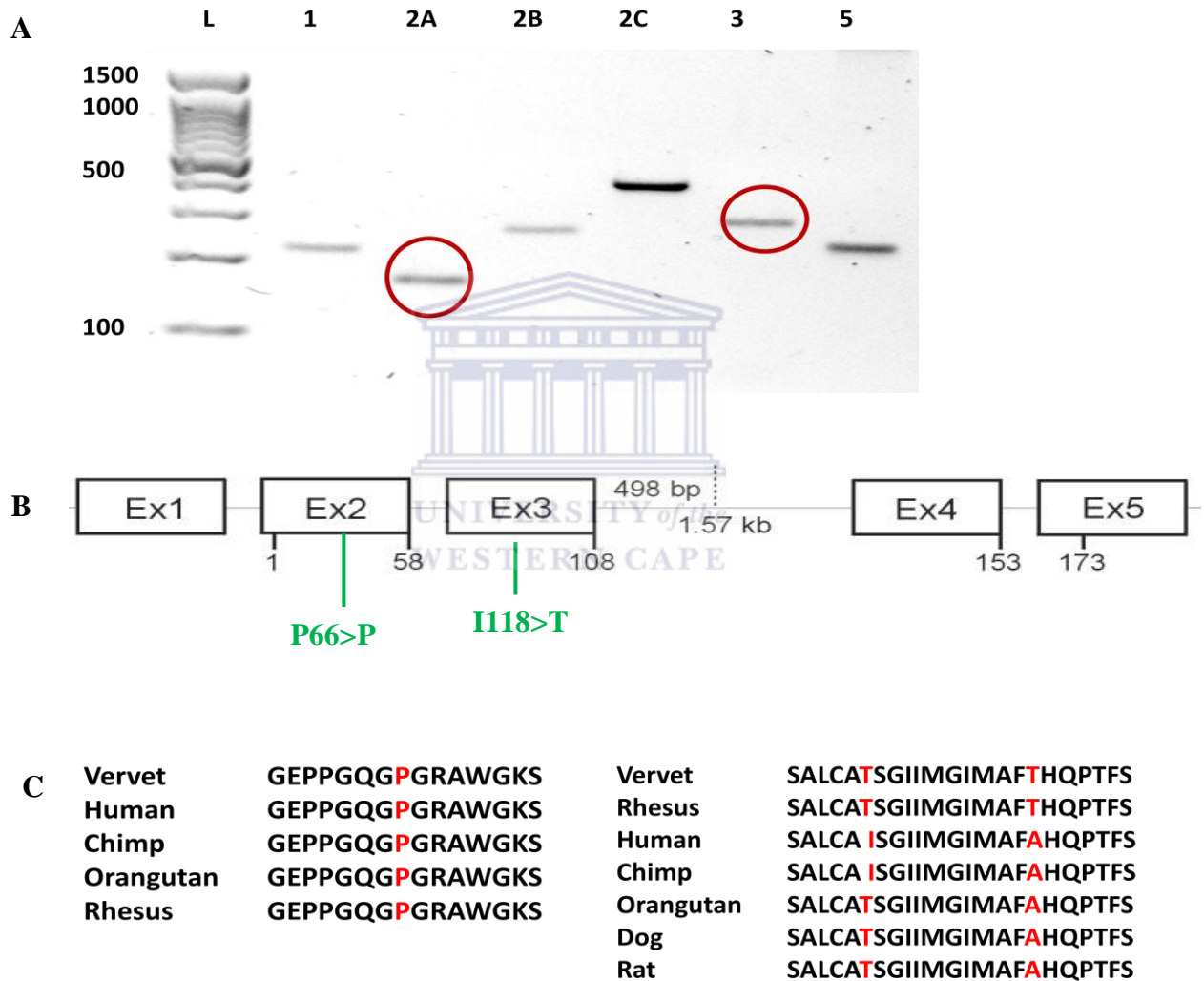


**Figure 3.27:** The structure of *GCNT2* with previously reported mutations. Adapted from Pras *et al.*, 2004. The mutations that were identified in Vervet monkeys are indicated in green.



### 3.2.4 DNA sequence variants in *LIM2*

Screening of five coding exons of *LIM2* (Figure 3.28A) revealed no mutations in all the cataract affected individuals. However, three single nucleotide polymorphisms (SNPs) were noted in exon 2 (P66>P) and exon 3 (I118>T and A127>T) (Figure 3.28B and C).



**Figure 3.28:** Genetic analysis of selected captive-bred vervet monkeys. **A)** A 2% agarose gel electrophoresis showing coding exons of *LIM2* gene with SNPs in exon 2 and 3 (red circles). **B)** The schematic structure of *LIM2* illustrating SNPs that were identified in Vervet monkeys (green) (Adapted from Shiels *et al.*, 2007). **C)** The protein sequence alignment among different species for exon 2 and 3, respectively.

## CHAPTER FOUR

### Discussion and Conclusion

#### 4.1 Introduction

Increasing incidents of cataract have been observed within a colony of captive-bred Vervet monkeys to the extent that it had a direct impact on the productivity. Both genders were involved and parents of the affected individuals were asymptomatic for the two observed cataract morphologies (Total and Y-sutural). This has become a colony-wide problem given that more cases are observed each year. Since parents of affected individuals are always free of cataract, a possibility of autosomal recessive transmission mode exists; excluding autosomal dominant and X-linked. The detailed family history and data reported by De Villiers *et al.*, (2001) assisted in classifying the cataract phenotypes. Clinically, eyes shared several common features like similar age of onset of the disease, bilateral or unilateral occurrence, similar clouding patterns, duration of clinical development and the lack of possible infectious agents. Independent factors such as microorganisms, diabetes mellitus and calcium concentrations that contribute to cataract development (Shiels and Hejtmancik, 2007) were taken into consideration for this study, however, since parents were asymptomatic, it was speculated that the main cause of cataract in the colony was genetic. This is supported by Plesker *et al.*, 2005; Kessler and Rawlins 1985 in similar studies which included the aforementioned factors on cataract affected nonhuman primates. For the purpose of this study, the above mentioned external factors were excluded.

Since cataract formation poses a direct health problem on the MRC captive bred Vervet monkey colony, a systemic strategy was implemented to manage breeding program. In order to minimize cataract incidents with regards to the concept of recessive transmission, family history of cataract-free individuals and mating selection programs were screened.

Based on previous studies in humans, four congenital cataract genes (*HSF4*, *CRYAA*, *GCNT2* and *LIM2*) were prioritised for this study. These genes are known to cause autosomal recessive congenital cataract in humans and have not been studied in Vervet monkeys. Therefore, this study provides the first report of DNA variations in this species, even though the lack of its sequence availability resulted in the utilization of human and rhesus sequences as a reference. This is one of the limitations for this study, which made screening of certain exonic regions of those prioritized genes challenging, hence exon seven and eight of *HSF4* and exon 4 of *LIM2* were not covered, though the same regions were successfully amplified with human and Rhesus DNA. Since some exonic regions could not amplify, this led to the conclusion that those specific regions within the two genes are missing in the Vervet sequence genome.

Twenty sequence variants were identified in this study. One sequence variant in *HSF4*, two in *CRYAA* and four in *GCNT2* were clearly disease-causing. Ten silent mutations and three SNPs were also identified, which were all located in conserved regions when aligned with other species. Since the mode of transmission is recessive, single parents of the affected individuals

(carriers) were included in the study. The sequence variants identified in cataract subjects and carriers were not present in the controls.

### ***HSF4***

According to literature, members of the heat shock family have three functional regions: (DBD, linker region and HR-A and B) (Figure 3.11) (Nakai *et al.*, 1997). One of the novel silent mutations reported in this study (R116R) (Figure 3.8) was found to be located in the DBD region, which is known to be important for binding trimers to heat shock elements (HSE). Mutations in this region are linked to AD congenital cataract (Enoki *et al.*, 2010). The R116R mutation was associated with unilateral and bilateral AR cataracts affecting both phenotypes in the Vervet monkey. This mutation was observed in three carriers where their offsprings were also affected (Appendix A2). In addition, two families that had a family member interacting with multiple partners in the E-cage shared this mutation (Figure 3.1 and 3.5). This could mean that the same individual was mated with different partners of which one of them was a carrier for R116R. The mutation occurred at a highly conserved arginine residue and was corresponding to the mutation that was reported by Shi *et al.*, 2008 (R116C), which is associated with age related cataract in patients that were 60-69 years old. However, this was not the case for R116R due to the fact that all the individuals were juveniles.



The effect of this silent mutation in Vervet monkeys is unknown. However, it can be speculated that it will exhibit the same changes as R116C since they are located in the same region together with other reported mutations (L115>P and R120>C) that are known to inhibit oligomerization of *HSF4b*, the binding of HSE to DNA and the transcription ability of *HSF4b* (Enoki *et al.*, 2010). In addition, silent mutations are known to interfere with signals for Ribonucleic acid (RNA) splicing, RNA folding, microRNA binding, regulation, translation rate and even protein folding (Cartegni *et al* 2002; Chamary *et al* 2006; Itzovitz and Alon 2007).

The other two identified mutations of *HSF4* (L245L and P421L) (Figure 3.9 and 3.10) are located in the distal end of HR where most of recessive cataract mutations are known to be located (Sajjad *et al.*, 2008). Sajjad *et al.*, (2008) reported a mutation (R405X) in the same region causing a premature termination associated with AR congenital cataract, and resulting in a complete loss of function of HSF4 protein in affected homozygotes. The possibility remains that slight changes in the transcriptional activity of HSF4 can result in a loss of its ability to oligomerize and bind to HSE, thus resulting in cataract formation (Mou *et al.*, 2010). The association of this gene with two different modes of inheritance can be explained by location and severity of the mutation (Sajjad *et al.*, 2008). For example, all known dominant mutations lie within DBD, whereas recessive mutations are more to the distal end of HR (Hejtmancik, 2008; Sajjad *et al.*, 2008). Based on the locations in which the mutations occurred in Vervet monkeys, it can be argued that recessive mutations are also located in the DBD region not just dominant mutations.

In addition, the mutations that are reported above (R116R, L245L and P421L) in Vervet monkeys were conserved in other species (Figure 3.8 – 3.10). The mechanisms by which they contribute to cataract phenotypes needs further investigation.

### **CRYAA**

Two missense mutations (S134W and K166>N) were identified in CRYAA. S134W is located in the ACD region (Figure 3.14) which is known to be essential for normal functioning of  $\alpha$  crystallines (Hansen *et al.*, 2007). According to literature, R116H mutation was reported by Gu *et al.*, (2008) to be located in this region. However, this mutation was linked to AD total cataract. S134W on the other hand was associated with unilateral and bilateral cataracts affecting both phenotypes (total and Y-sutural) of the Vervet monkey. Since both mutations are located in the same ACD region and R116H has already been reported to interfere with chaperone activity within  $\alpha$  crystallines, it can be speculated that S134W may be driven by similar mechanisms with similar effects despite being linked to AR instead of AD.

The second mutation K166>N is located in the C-terminal region (Figure 3.14). So far there are no reported mutations in the C-terminal region which are suggested to play a role in oligomerization, however, the K166>N in Vervet monkeys caused a change in amino acid polarity (basic to neutral). This might alter the normal functioning of the protein product.

## ***GCNT2***

Screening of three *GCNT2* transcripts resulted in identification of eleven sequence variations. Only four missense mutations (M258>V, V16>I, Y122>F and S24>N) were identified in this gene, the rest were silent mutations (appendix B, Table B1). The M258>V was identified in exon one of transcript A. So far, no mutations have been reported in this exon, this mutation is reported for the first time in this region.

The other two mutations (V16>I, Y122>F) were identified in exon one of transcript B and according to Borck *et al.*, (2012), mutations in transcript 1B play an important role in cataract development. These two sequence variants are the first missense mutations to be identified in exon one of this transcript since previously reported mutations were deletions (Yu *et al.*, 2001; Borck *et al.*, 2012). It was interesting to note both mutations are linked to bilateral cataracts in the Vervet monkeys. Since this is the first report, the function and mechanism by which these mutations may cause cataracts is unknown.

The fourth mutation (S24>N) was identified in exon one of transcript C. Mutations in this transcript are associated with adult i phenotype (Yu *et al.*, 2003; Pras *et al.*, 2004; Wussuki-Lior *et al.*, 2011; Borck *et al.*, 2012). Previously reported mutations (A169>T and R228Q) in exon one were only found in patients with adult i phenotype. There is an association between congenital cataract and adult i phenotype (Yu *et al.*, 2003; Pras *et al.*, 2004; Wussuki-Lior *et al.*, 2011; Borck *et al.*, 2012). However, this association need to be investigated further in Vervet monkeys.

## ***LIM2***

There were no mutations identified in *LIM2*, However three SNPs were identified in exon 2 (P66>P) and exon 3 (I118>T and A127>T) (Figure 3.28B). In human studies, only two mutations linked to recessive cataract have been identified (F105>V and G154>E) thus far (Pras *et al.*, 2002; Ponnam *et al.*, 2008). Even though *LIM2* is generally considered to be the second most abundant lens fibre cell protein, little is known about its function. Based on our findings and Pras *et al.*, 2002 and Ponnam *et al.*, 2008, it is clear that this gene is not prone to mutations that may affect cataracts.

### **4.2 Limitations of the study.**

One of the main limitations of this study was the unavailability of Vervet monkey genome sequence: however recent progress has been made to include the first draft of the Vervet monkey sequence in the GenBank database. Since the genome sequence of Vervet monkeys was not yet available, designing primers for the above mentioned genes was challenging, therefore, human and rhesus sequences were used as a reference.

### **4.3 Strengths of the study**

Nonhuman primates are closely related to humans, both in evolutionary and genetic terms, therefore; results obtained from this study can contribute to the existing knowledge about factors influencing human visual disorders. The study provides more insight into inherited cataract by using molecular techniques to fill in the gap that exist in literature on cataracts that developed in

animals. It will also contribute to a better understanding of the biology and physiology of nonhuman primates. Furthermore, it is the first time genetic aspects have been defined in the Vervet monkey to address one of the diseases of the eye affecting both humans and animals. Results obtained will also contribute to defining the molecular genetics of the Vervet monkey especially in the light that its genome sequence has not yet been fully mapped.

#### **4.4 Conclusion**

The main aim for this study was to identify sequence variants within the Vervet colony, to facilitate a strategy to manage breeding programs and minimize cataract cases. Two types of cataract phenotypes (total and Y-sutural) were observed and all affected monkeys and their nonsymptomatic parents were selected for DNA screening. Four genes (*HSF4*, *CRYAA*, *GCNT2* and *LIM2*) known to be responsible for autosomal recessive congenital cataract in humans were prioritized. These genes were fully screened, sequenced and analysed. Twenty sequence variations were identified affecting both phenotypes; however, two mutations in *GCNT2* were only identified in total cataract individuals. Only one variation was located in a region that was previously reported in humans, the remaining mutations are reported for the first time in Vervet monkeys, therefore novel. Most of the sequence variants that were shared between the two phenotypes were observed in individuals within the same families. Based on these findings, it can be concluded that the three candidate genes (*HSF4*, *CRYAA* and *GCNT2*) harbour mutations that are responsible for both phenotypes.

Ever since the study commenced, the family members of the cataract affected monkeys were excluded from the breeding programs based on their family history and this lowered the cataract cases in the colony. This means that our objectives have been met; however, the mechanism by which these selected genes results in congenital cataract must be further investigated in order to determine the effects of these reported mutations in respect to the gene functionality. To strengthen the results, it will be beneficial to screen the entire colony for mutations since this is inherited in an autosomal recessive manner.



## REFERENCES

Amaya L, Taylor D, Russwll-Eggitt I, Nischal K.K and Lengyel D (2003). **The morphology and natural history of childhood cataracts.** *Surv Ophthalmol*, 48 (2): 125-144

Andley U.P (2007). **Crystallins in the eye: function and pathology.** *Prog Retin Eye Res*, 26: 78–98.

Azuma N, Hirakiyama A, Inoue T, Asaka A and Yamada M (2000) **Mutations of a human homologue of the *Drosophila* eyes absent gene (EYA1) detected in patients with congenital cataracts and ocular anterior segment anomalies.** *Hum Mol Genet*, 9:363-6.

Bagchi M, Katar M and Maisel H (2002). **Heat shock of adult and embryonic human ocular lenses.** *J Cel Bio*, 84: 278-284.

Berry V, Francis P, Reddy MA, Collyer D, Vithana E, MacKay I, Dawson G, Carey AH, Moore A, Bhattacharya SS, Quinlan RA (2001). Alpha-B crystallin gene (CRYAB) mutation causes dominant congenital posterior polar cataract in humans. *Am J Hum Genet*, 69:1141–1145

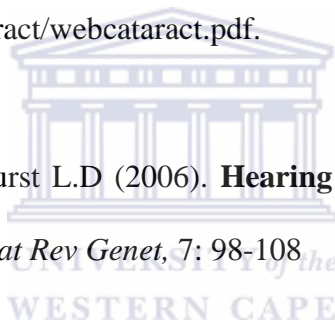
Borck G, Kakar N, Hoch J, friedrich K, Freudenberg J, Nurnberg P, Oldenburg J, Ahmad J and Kubisch C (2012). **An Alu repeat-mediated genomic GCNT2 deletion underlies congenital cataracts and adult i blood group.** *Hum Genet*, 131: 209-216.

Bu L, Jin Y, Shi Y, Chu R, Ban A, Eiberg H, Andres L, Jiang H, Zheng G, Qian M, Cui B, Xia Y, Liu J, Hu L, Zhao G, Hayden M.R and Kong X (2002). **Mutant DNA-binding domain of HSF4 is associated with autosomal dominant lamellar and Marner cataract.** *Nat Genet*, 31(3): 276-8.

Cartegni L, Chew SL, Krainer A.R (2002). **Listening to silence and understanding nonsense: exonic mutations that affect splicing.** *Nat Rev Genet*, 3: 285-98.

Cataract: **What you should know.** National Eye Institute (2010).  
<http://www.nei.nih.gov/health/cataract/webcataract.pdf>.

Chamary J.V, Parmley J.L and Hurst L.D (2006). **Hearing silence: non-neutral evolution at synonymous sites in mammals.** *Nat Rev Genet*, 7: 98-108



Chen j, Ma Z, Jiao X, Fariss R, Kantorow W.L, Kantorow M, Pras E, Frydman M, Pras E, Riazuddin S, Riazuddin S.A and Hejtmancik J.F (2011). **Mutations in FYCO1 cause autosomal-recessive congenital cataracts.** *Am J Hum Genet*, 88: 827-838.

Clark A.R, Lubsen NH and Slingsby C (2012). **sHSP in the eye lens: Crystallin mutations, cataract and proteostasis.** *Int J Biochem Cell Biol*, 44; 1687-1697.

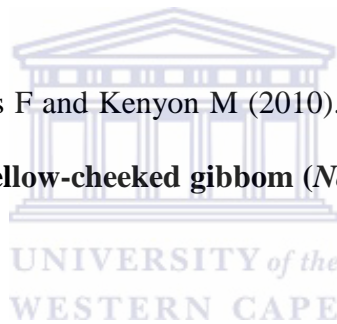


Devi R.R, Yao W, Vijayalakshmi P, Sergeev Y.V, Sundaresan P, Hejtmancik J.F (2008). **Crystallin gene mutations in Indian families with inherited pediatric cataract.** *Mol Vis*, 14: 1157–1170.

MacLaren N (2008). **Monkey's cataract removed.** *Deseret News*, 30 April.

De Villiers C, Seier J.V and Dhansay M.A (2001). **Probable genetic origin for a large number of cataracts among captive-bred vervet monkeys (*Chlorocebus aethiops*).** *Am J Primatol*, 55:43-48.

Duy T.P, Hoang T.H, van den Bos F and Kenyon M (2010). **Successful cataract removal and lens replacement on a rescued yellow-cheeked gibbon (*Nomascus gabriellae*).** *Vietnamese J. Primatol*, 4:69-74.



Eichler E.E and DeJong P.J (2002). **Biomedical applications and studies of molecular evolution: A proposal for a primate genomic library resource.** *Genome Research*, 12: 673-678

Enoki Y, Mukoda Y, Furutani C, and Sakurai H (2010). **DNA-binding and transcriptional activities of human HSF4 containing mutations that associate with congenital and age-related cataracts.** *Molecular Basis of Disease*, 1802(9): 749-53.

Foster A (1999). **Cataract—A global perspective: output, outcome and outlay.** *Eye*, 13 449–453.

Forsheew T, Johnson C.A, Khaliq S, Pasha S, Willis C, Abbasi R, Tee L, Ursula Smith U, Trembath RC, Mehdi SQ, Moore A.T, Maher ER (2005). **Locus heterogeneity in autosomal recessive congenital cataracts:linkage to 9q and germline HSF4 mutations.** *Hum Genet*, 117: 452–459.

Francis P.J, Berry V, Bhattacharya S.S and Moore A.T (2000). **The genetic of childhood cataract.** *J Med Genet*, 37: 481-488.

Francis P.J, Berry V, Moore A.T, and Bhattacharya S.S (1999). **Lens Biology: development and human cataractogenesis.** *Trends genet*, 15: 191-196.

Frejtag W, Zhang Y, Dai R, Anderson M.G and Mivechi N.F (2001): **Heat shockfactor-4 (HSF-4a) represses basal transcription through interaction with TFIIF.** *J Biol Chem*, 276: 14685-14694.

Fujimoto M, Izu H, Seki K, Fukuda K, Nishida T, Yamada S, Kato K, Yonemura S, Inouye S and Kakai A (2004). **HSF4 is required for normal cell growth and differentiation during mouse lens development.** *EMBO Journal*, 23: 4297-4306.

GeneCards (2012) available online (<http://www.genecards.org/>)

Gill D, Klose R, Munier F.L, McFadden M, Priston M, Billingsley G, Ducrey N, Schorderet D.F and Heon E (2000). **Genetic heterogeneity of the Coppock-like cataract: a mutation in CRYBB2 on chromosome 22q11.2.** *Invest Ophthalmol Vis Sci*, 41: 159-165.

Gonen T, Grey A.C, Jacobs M.D, Donaldson P.J and Kistler J. (2001). **MP20, the second most abundant lens membrane protein and member of the tetraspanin superfamily, joins the list of ligands of galectin-3.** *BMC Cell Biol*, 2:17.

Graw J (2009). **Genetics of crystallins: cataract and beyond.** *Exp Eye Res*, 88: 173-189.

Graw J, Löster J, Soewarto D, Fuchs H, Meyer B, Reis A, Wolf E, Balling R, and Hrabé de Angelis M (2001). **Characterization of a New Dominant V124E Mutation in the Mouse  $\alpha$ -Crystallin-Encoding Gene.** *Investigative Ophthalmology & Visual Science*, 42 (12): 2909-15.

Gu F, Luo W, Li X, Wang Z, Lu S, Zhang M, Zhao B, Zhu S, Feng S, Yan Y, Huang S and Ma X (2008). **A novel mutation in  $\alpha$ A-crystallin (CRYAA) caused autosomal dominant congenital cataract in a large Chinese family.** *Hum Mut*, 29(5): 1-8

Hansen L, Yao W, Eiberg H, Kjaer K.W, baggesen K, Hejtmancik J.F and Rosenberg T (2007). **Genetic Heterogeneity in Microcornea-Cataract: Five Novel Mutations in CRYAA, CRYGD, and GJA8.** *Invest Ophthalmol Vis Sci*, 48 (9): 3937-3944.

Hejtmancik J.F (2008). **Congenital cataract and their molecular genetics**. *Cell & Dev Bio*, 19: 134-149

Hejtmancik J.F (1998). **The genetics of cataract: our vision becomes clearer (editorial)**. *Am J Hum Genet*, 62: 520–525.

Hejtmancik J.F, Kaiser-Kupfer MI, Piatigorsky J (2001) **Molecular biology and inherited disorders of the eye lens**. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) *The metabolic basis of inherited disease*, vol 7. McGraw Hill, New York, pp 4325–4349.

Hejtmancik J.F and Smaoui N (2003). **Molecular genetics of cataract**. In: Wissinger B, Kohl S, Langenbeck U, eds. *Genetics in Ophthalmology*. Basel, Switzerland: S Karger: 67-82.

Heon E, Liu S, Billingsley G, Bernasconi O, Tsiflidis C, Schorderet D.F, Munier F.L and Tsiflidis C (1998). **Gene localization for aculeiform cataract, on chromosome 2q33-35**. *Am J Hum Genet*, 63: 921-6.

Horwitz J (1992) **Alpha-crystallin can function as a molecular chaperone**. *Proc Natl Acad Sci*, 89:10449-10453.

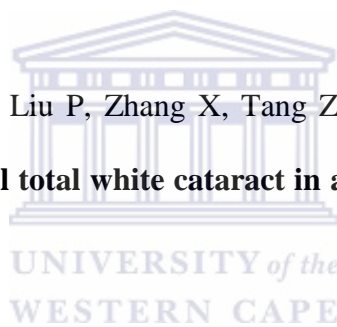
Huang B and He W (2010). **Molecular characteristics of inherited congenital cataracts**. *Euro J Med Genet*, 53: 347-357.

Ionides A, Francis P, Berry V, Mackay D, Bhattacharya S and Shiels A (1999). **Clinical and genetic heterogeneity in autosomal dominant congenital cataract.** *Br J Ophthalmol*, 83:802-808.

Izkovitz S and Alon U (2007). **The genetic code is nearly optimal for allowing additional information within protein-coding sequences.** *Genome Res*, 17: 405-12

Kannabiran C and Balasubramanian D (2000). **Molecular genetics of cataracts.** *Indian J Ophthalmol*, 48: 5-13

Ke T, Wang Q.K, Ji B, Wang X, Liu P, Zhang X, Tang Z, Ren X and Liu M (2006). **Novel HSF4 mutation causes congenital total white cataract in a Chinese family.** *Am J Ophthalmol*, 142: 298-303.



Kessler M.J and Rawlins R.G (1985). **Congenital cataract in a Free-Ranging Rhesus Monkey.** *J Med Primatol*, 14:225-228.

Khan A.O Aldahmesh M.A and Meyer B (2007). **Recessive congenital total cataract with microcornea and heterozygote carrier signs caused by a novel missense CRYAA mutation (R54C).** *Am J Ophthalmol*, 144: 949-952.

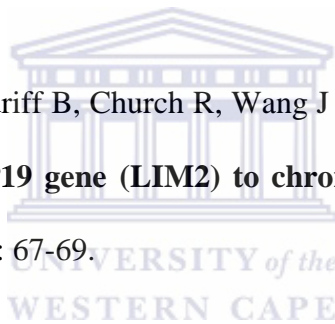
King F.A, Yarbrough C.J, Anderson D.C, Gordon T.P, Gould K.G (1988). **Primates.** *Science*, 240:1475-81.

Klopp N, Heon E, Billingsley G, Illig T, Wjst M, Rudolph G and Graw J (2003). **Further genetic heterogeneity for autosomal dominant human sutural cataracts.** *Ophthalmic Res*, 35: 71-7.

Kumar L.V.S, Ramakrishna T and Rao C.M (1999). **Structural and functional consequences of the mutation of a conserved arginine residue in alpha-A and alpha-B crystallins.** *J Biol Chem*, 274: 24137–24141.

Lambert S.R and Drack A.V (1996). **Infantile cataracts.** *Surv Ophthalmol*, 40(6): 427-458

Lieuallen K, Christensen M, Brandriff B, Church R, Wang J and Lennon G (1994). **Assignment of the human lens fiber cell MP19 gene (LIM2) to chromosome 19q13.4 and adjacent to ETFB, Somat.** *Cell Mol Genet*, 20: 67-69.



Louis C.F, Hogan P, Visco L and Strasburg G (1990). **Identity of the calmodulin-binding proteins in bovine lens plasma membranes.** *Exp Eye Res*, 50: 495–503.

Maniatis T, Fritsch F, Sambrook J. (1989) **Molecular cloning: a laboratory manual.** Cold Spring Harbour Laboratory, Cold Spring Harbour, New York.

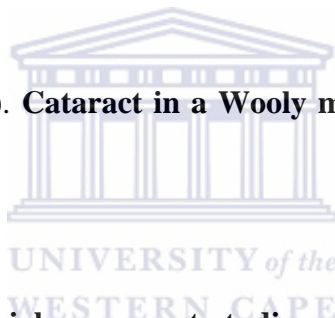
Mou L, Xu J, Li W, Lei X, Wu Y, Xu G, Kong X and Xu G (2010). **Identification of Vimentin as a Novel Target of HSF4 in Lens Development and Cataract by Proteomic Analysis.** *Ophthalmol Vis Sci*, 51: 396-404.

Nakai A, Tanabe M, Kawazoe Y, Inazawa J, Morimoto R.I and Nagata K (1997). **HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator.** *Mol Cell Biol*, 17: 469-481.

Ollivet F, Beltran W, Lecu A and Chahory S (2000). **Ocular findings in a colony of mouse lemurs (*Microcebus murinus*) from the Paris Zoo.** In: proceedings of the American association of zoo veterinarians. Baer CK, Willette MM(eds). Orlando, FL: Conference 148-150.

Ophthalmological Society of South Africa (2007). Available ([http// www.ossa.co.za](http://www.ossa.co.za)).

Peiffer R.L and Gelatt K.N (1976). **Cataract in a Woolly monkey.** *Mod Vet Pract*, 57(8):609-610.



Piatigorsky J (1990). **Molecular biology: recent studies on enzyme/ crystallins and alpha-crystallin gene expression.** *Exp Eye Res*, 50: 725–8.

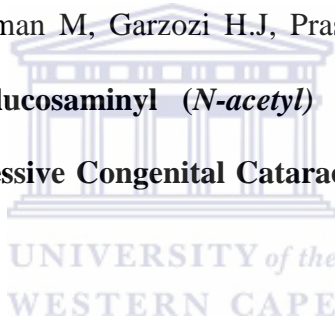
Plesker R, Hetzel U and Schmidt W (2005). **Cataract in laboratory colony of African green monkeys (*Chlorocebus aethiops*).** *J Med Primatol*, 34:139-146.

Ponnam S.P.G, Ramesha K, Tejawani S, Matalia J, and Kannabiran C (2008). **A Missence Mutation in *LIM 2* Causes Autosomal Recessive Congenital Cataract.** *Molecular Vision*, 14: 1204-1208.

Pras E, Frydman M and Levy-Nissenbaum E, Bakhan T, Raz J, Assia E.I, Goldman B and Pras E.A (2000). **Nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family.** *Invest Ophthalmol Vis Sci*, 41 (11): 3511-3515.

Pras E, Levy-Nissenbaum E, Bakhan T, Lahat H, Assia E, Geffen- Carmi N, Frydman M, Goldman B and Pras E.A (2002) **Missense mutation in the LIM2 gene is associated with autosomal recessive presenile cataract in an inbred Iraqi Jewish family.** *Am J Hum Genet*, 70(5): 1363–1367.

Pras E, Raz J, Yabalom V, Frydman M, Garzozzi H.J, Pras E and Hejtmancik J.F (2004). **A Nonsense Mutation in the Glucosaminyl (*N*-acetyl) Transferase 2 Gene (*GCNT2*): Association with Autosomal Recessive Congenital Cataracts.** *Invest Ophth & Vis Sci*, 45 (6): 1940-45.



Rees M.I, Watts P, Fenton I, Clarke A, Snell R.G, Owen M.J, Gray J (2000). **Further evidence of autosomal dominant congenital zonular pulverulent cataracts linked to 13q11 (CZP3) and a novel mutation in connexin 46 (GJA3).** *Hum Genet*, 106(2):206-9.

Reddy M.A, Francis P.J, Berry V, Bhattacharya S.S and Moore A.T (2004). **Molecular genetic basis of inherited cataract and associated phenotypes.** *Surv Ophthalmol*, 49: 300-315.



Rogaev E, Korovaitseva G, Farrer L, Petrin A and Keryanov S (1996). **Linkage of polymorphic congenital cataract to the gamma-crystallin locus on human chromosome 2q33-35.** *Hum Mol Genet*, **5**:699-703.

Rogers J, Garcia R, Shelledy W, Kaplan J, Arya A, Johnson Z, Bergstrom M, Novakowski L, Nair P, Vinson A, Newman D, Heckman G and Cameron J (2006). **An initial genetic map of the rhesus macaque (*Macaca mulatta*) genome using human microsatellite loci.** *Genomics*, **87**: 30-38.

Sajjad N, Goebel I, Kakar N, Cheema A.M, Kubisch C and Ahmad J (2008). **A novel gene mutation (p.R405X) causing autosomal recessive congenital cataract in a large consanguineous family from Pakistan.** *BMC Med Genet*, **9**: 99.

Sasaki Y, Kodama r, Iwashige S, Fujishima J, Yoshikawa T, Kamimura Y and Maeda H (2011). **Bilateral cataract in Cynomolgus monkey.** *J Toxicol Pathol*, **24**: 69-73.

Seier J.V, Chauke C.G and Van Heerden J.J (2012). **Standard Operating Procedures.** Unpublished MRC Primate Unit Report. Medical Research Council.

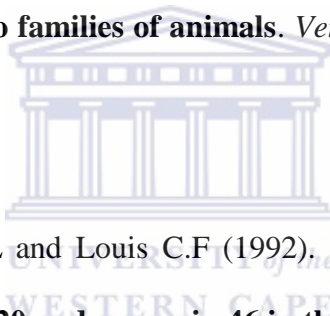
Shi Y, Shi X, Jin Y, Miao A, Bu L, He J, Jiang H, Lu Y, Kong X, Hu L (2008). **Mutation Screening of *HSF4* in 150 age-related cataract patients.** *Molecular Vision*, **14**:1850-1855.

Shiels A and Hejtmancik J.F (2007). **Genetic origins of cataract.** *Arch Ophthalmol*, 125:165–73.

Smaoui N, Beltaief O, BenHamed S, M'Rad R, Maazoul F, Ouertani A, Chaabouni H, and Hejtmancik J.F (2004). **A homozygous splice mutation in the HSF4 gene is associated with an autosomal recessive congenital cataract.** *Invest Ophthalmol Vis Sci*, 45: 2716-21.

Song S, Landsbury A, Dahm R, Liu Y, Zhang Q and Quinlan R.A (1997). **Functions of the intermediate filament cytoskeleton in the eye lens.** *J. Clin. Invest*, 119:1837–1848.

Souri E (1973). **Eye disease in two families of animals.** *Vet Med Small Animal Clin*, 68: 1011-1014.



Tenbroek E, Arneson M, Jarvis L and Louis C.F (1992). **The distribution of the fiber cell intrinsic membrane proteins MP20 and connexin 46 in the bovine lens.** *J Cell Sci*, 103: 245–57.

TenBroek E.M, Johnson R and Louis C.F (1994). **Cell-to-cell communication in a differentiating ovine lens culture system.** *Invest Ophthalmol Vis Sci*, 35: 215–28.

Vanita, Sarhadi V, Reis A, Jung M, Singh D, Sperling K, Singh JR, Bürger J (2001). **A unique form of autosomal dominant cataract explained by gene conversion between beta-crystallin B2 and its pseudogene.** *J Med Genet*, 38 (6): 392-6.

VandeBerg J.L (1995). **Genetics of nonhuman primates**. In: Bennett BT, Abee CR, Henrickson R (eds) *Nonhuman primates in biomedical research: biology and management*. Academic, San Diego, pp 129–146.

VandeBerg J.L and Williams-Blangero S (1997). **Advantages and limitations of nonhuman primates as animal models in genetic research on complex diseases**. *J Med Primatol*, 26: 113-119.

Wussuki-Lior O, Abu-Horowitz A, Netzer I, Almer Z, Morad Y, Goldich Y, Yahalom V, Pras E.I and Pras E.R (2011). **Hematologic biomarkers in childhood cataracts**. *Mol Vis*, 17: 1011-1015.

Wolf N.S, Li Y, Pendergrass W, Schneider C and Turturro A (2000). **Normal Mouse and rat strains as models for age related cataract and the effect of caloric restriction on its development**. *Exp Eye Res*, 70: 683-692.

Yamada K, Tomita H, Yoshiura K, Kondo S, Wakui K, Fukushima Y, Ikegawa S, Nakumura Y, Amamiya T and Niikawa N (2000). **An autosomal dominant posterior polar cataract locus maps to human chromosome 20p12-q12**. *Eur J Hum Genet*, 8 (7): 535-539.

Yu L.C, Twu Y.C and Chou M.L, Reid M.E, Gray A.R, Moulds J.M, Chang C and Lin M (2003). **The molecular genetics of the human I locus and molecular background explain the partial association of adult I phenotype with congenital cataracts**. *Blood*, 101: 2081-88.

Zhang Q, Guo X, Xiao X, Yi J, Jia X and Hejtmancik J.F (2004). **Clinical description and genome wide linkage study of Y-sutural cataract and myopia in a Chinese family.** *Molecular Vis*, 10: 890-900.

Zhang R.L, Samuelsn D.A, Zhang Z.G, Reddy V.N and Shastry B.S (1991). **Analysis of eye lens specific genes in congenital hereditary cataracts and microphthalmia of the miniature schnauzer dog.** *Invest ophthalmol Vis Scie*, 32 (9): 2662-2665.



## APPENDIX A

### 1. MATERIALS AND SUPPLIERS

Agarose	Whitehead Scientific
Boric acid	Merck
EDTA	BDH
EDTA blood collection tubes	Pathcare
Ethanol	BDH
Paxgene Blood collection tubes	Qiagen
PCR Master mix	Promega
Tris-base	Promega
Wizard SV Gel and PCR Clean-Up System	Promega
100bp DNA ladder	Biocom Biotech
Loading gel	Promega
Ethidium bromide	Promega
Nuclease free water	Promega
Whole blood separation kit	Separations



Oligonucleotides/primers

UCT

Sequencing Ready Reaction Kit

Applied Biosystems

## **2. Extraction of DNA from whole blood**

Five hundred microliter (500  $\mu$ l) of buffer AP1 was added into a 1.5ml microfuge tube. A 200-250  $\mu$ l of anti-coagulated blood was added. The cap of the microfuge tube was closed and mixed by vortexing at top speed for 10 seconds. Buffer AP2 (100  $\mu$ l) was added and mixed by vortexing at top speed for 10 seconds. The mixture was centrifuged at 12,000 x g for 10 minutes at ambient temperature to pellet cellular debris. A miniprep column was placed into a 2ml Microfuge tube. The clarified supernatant that was obtained from centrifugation was pipette into the Miniprep column and centrifuged at 6,000 X g for 1 minute. The filtrate from the 2ml microfuge tube was discarded. The miniprep was placed back into the 2ml microfuge tube. Buffer W1A (700  $\mu$ l) was pipette into the miniprep column and allowed to stand at room temperature for 2 minutes and centrifuged at 6,000 X g for 1 minutes. The filtrate from the 2ml microfuge tube was discarded. The miniprep was placed back into the 2ml microfuge tube. Buffer W2 (800  $\mu$ l) was pipette into the miniprep column and allowed to stand at room temperature for 2 minutes and centrifuged at 12,000 X g for 1 minutes. The filtrate from the 2ml microfuge tube was discarded. The miniprep was placed back into the 2ml microfuge tube. Buffer W2 (500  $\mu$ l) was pipette into the miniprep column and allowed to stand at room temperature for 2 minutes and centrifuged at 12,000 X g for 1 minutes. The filtrate from the 2ml microfuge tube was discarded. The miniprep was placed back into the 2ml microfuge tube and centrifuged at 12,000 X g for 1 minute. The miniprep column was placed into a 1.5ml microfuge

tube. Buffer TE (80-200  $\mu$ l) was added and allowed to stand at room temperature for 1 minute. The mixture was then centrifuged at 12, 000 X g for 1 minute to elute genomic DNA.

### 3. Primers for all the “candidate genes”

**Table A1:** The list of designed primers based on Human and Rhesus macaque sequences.

GENE	EXON	STRAND	SEQUENCE	LENGTH (bp)	T <sub>m</sub>
<i>CRYAA</i>	1	Forward	CCTTAATGCCTCCATTCTGC	408	58.4
		Reverse	GACGGAGCAAGACCAGAGTC		
	2	Forward	ACGTTTGGATTTCAGGTTTCG	240	55
		Reverse	AAGGCATGGTGCAGGTGT		
	3	Forward	ACATTTCCCGTGAGTCCAC	749	55
		Reverse	ATGGAGACAGCACCAGCAG		
<i>HSF4</i>	4 to 5	Forward	GGCGGCGTTCTTGGTAGAGCGG	419	55.9
		Reverse	GGACTGGGTCGCAGGAGCAAG		
	6	Forward	ATGAGCAAAGAGGAGGAGGGGTG	504	55.9
		Reverse	CGTGGCTCTGCCAAGTGTC		
	7 to 8	Forward	CCCAGCCTCGCCATTCTGTG	480	-
		Reverse	TTCCCGGTGAAGGAGTTTCCA		
	9	Forward	ACACAGGTCCCTGATGCTGGATG	164	58.4
		Reverse	AGGCTCTCCATAAGCCCAGCCAT		
	10	Forward	GTTCTGGCTCTCCCTGTGCCTAC	192	60.9
		Reverse	TCCCCCTTACCTCCTGCCATCA		
	11	Forward	TGGTTGAAGCTTTTCTCTGGTGCA	313	59.9
		Reverse	TGTGGGCTGGTAAGGGCTGTT		
	12	Forward	GCCAAAAGCAGTTCTGTCTGCAC	179	59.3
		Reverse	AGACCCACCAGGTCTCATGC		
	13	Forward	GGCACCCTGACCCAGAGCTC	147	60
		Reverse	GAGGGCTTGACTCAGCCACCC		

	14	Forward	CGGTTCTCACGCAGATGCAGCC		136	59.3
		Reverse	AGCTCAGCCCAATCACGGCGT			
	15	Forward	CATTGGGCGAGAGTGGGGAGGTTAA		416	64.3
		Reverse	GTCGGGGTAGTGGAGAGAGGCC			
<b><i>GCNT2</i></b>	1A(1)	Forward	TGTAGACACAGGTTGCAG	GTTAGCAGCA	508	53
		Reverse	GGTAGCTTCATCAAGGGTA			
	1A(2)	Forward	TAGCAGAAGCCTGTCATCAG		500	53
		Reverse	CCTTCAGATACTGAACTATTT C			
	1A(3)	Forward	AACACCTGCGGGCAAGACTT		570	55
		Reverse	CTTTTGTCCCTGTGAACAGAGCGGT T			
	1B(1)	Forward	AGACTTACAGATTGTGACGGT		411	56.9
		Reverse	TAGATATTTTGGGGCATGTA			
	1B(2)	Forward	CCATCATCACTTTGACACCT		429	54.3
		Reverse	CTTATCACATATGAAAGCTCT			
	1B(3)	Forward	CTCATGCAATTGGACGGACT		380	54.3
		Reverse	GAGTGAGAACTACATATATTCATTCCG GTT			
	1C(1)	Forward	GTA AATTCAGCCTCTCACACCAATC		437	48.4
		Reverse	GGGGCATATAGATAGCCCTGA			
	1C(2)	Forward	TGTCACGGTCATCCATAAAG		407	50.4
		Reverse	CTTGGTGGACATATTTAGTT			
	1C(3)	Forward	AGGATTTAAAGGGAAAAATATC		374	48.4
		Reverse	TGAGTCAGTTCTCTAGGTGAGCAG			
	2	Forward	ACATTGCAGGTGTTCCCTGGCTC		398	61
		Reverse	GAACGAGAGCCTCACCGTGGC			
	3	Forward	AGTTGTAGTTAGTCGGAGAGTACCT		494	61
		Reverse	TATAATTACGTAGCCAGGTCCTGAA			
<b><i>Lim2</i></b>	1	Forward	CCATTGTGTAGGGAGGCTTA		213	55
		Reverse	AGGTCCTGGGAGAAGAAGG			
	2A	Forward	CAG TTC CCT TCC TCCATCAAGTCC		159	53
		Reverse	ACTGCATCCAATGGTCTGTT			



2B	Forward	TGTACAGCTTCATGGGTGGT	255	52.7
	Reverse	TGGAATACAGGTGTCCTTGG		
2C	Forward	TACCTGCAGACAGACAGCAT	238	51
	Reverse	CCCAACTTAACCTTCAAACC		
3	Forward	TCATCTCAGAGGTAGCAGGCA	279	55
	Reverse	ATTGGGGTTTGAGATGAGAG		
4	Forward	AAAATCACACCCAGCCTTAG	248	-
	Reverse	ACTCTATCTGCTGCCCACTC		
5	Forward	GGTGTGGGGCTCTCTTG	231	51
	Reverse	CTAGGAACCAGGATTCA		

#### 4. Standard PCR procedures

**Table A2:** PCR sample preparation



Components (Promega)	Experiment (µl)	Control (µl)	Final concentration
<b>Master mix 2X</b>	12.5µl	12.5µl	1X
<b>Primers (R&amp;F) 10µM</b>	1.25µl x2	1.25µl x2	0.5µM
<b>DNA</b>	1µl	0µl	50ng
<b>Nuclease free water to</b>	25µl	25µl	N/A

**Table A3: PCR conditions**

<b>STEP</b>	<b>TEMPERATURE (°C)</b>	<b>TIME</b>	
<b>Initial denaturation</b>	95	5min	
<b>Final denaturation</b>	95	30sec	} <b>30 cycles</b>
<b>Annealing</b>	Primer dependant	30sec	
<b>Elongation</b>	72	1min	
<b>Final elongation</b>	72	5min	

### **5. 5X TBE buffer**

54g TBE

27.5 Boric acid

3.7422 EDTA

Add distilled water to a total volume of 1000ml.

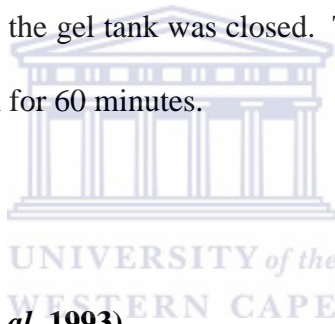


### **6. 1X TBE buffer**

100ml of 5X TBE was mixed with 400ml of distilled water to make a total volume of 500ml. this mixture was used to make and run agarose gel.

## 7. Preparation of agarose gel (1 or 2%)

One gram (1g) or 1.5g of agarose gel was weighed. Fifty millilitres (50ml or 75ml) of TBE working stock solution was added and mixed properly in a volumetric flask. The mixture was heated using microwave for 2minutes to dissolve the agarose. The mixture was cooled at room temperature. Ethidium bromide (1 or 2 $\mu$ l) was added to the mixture and swirled to mix. The combs were inserted into the gel tank properly. The gel was poured slowly into the tank. Using a disposable tip, bubbles were pushed away to the sides. The gel was left to set for 15 minutes or preferably 30min. The first lane of the gel was loaded with the marker. 10 $\mu$ l of prepared sample was pipette out and mixed with 2  $\mu$ l of 6x loading dye. The mixture was then loaded into each well. After loading all the samples, the gel tank was closed. The power- source was switched on and the gel was runned at 100V/cm for 60 minutes.



## 8. Tracking Dye III (Maniatis *et al*, 1993)

0.25% bromophenol blue

0.25% xylene cyanol FF

30% glycerol in ddH<sub>2</sub>O

Store at 4°C

## 9. Protocol for purification of PCR product

Following electrophoresis, the desired band was cut out of the electrophoresis gel and its weight was determined. Membrane Binding Solution (10  $\mu$ l) was added for every 10 mg of the gel slice. The mixture was vortexed and incubated at 50- 65°C until the gel was dissolved. An equal volume of membrane binding solution was added to the PCR product. The SV Minicolumn was inserted into collection tube. A dissolved gel mixture or prepared PCR product was transferred to the Minicolumn assembly. The mixture was incubated at room temperature for 1 minute. The mixture was centrifuged at 16, 000 X g for 1 minute. A flow through was discarded and Minicolumn was inserted into collection tube. Membranes wash solution (700  $\mu$ l) and centrifuged at 1 minute. Flow through was discarded and Minicolumn was reinserted into collection tube. The washing step was repeated with 500  $\mu$ l membrane wash solution. The mixture was centrifuged at 16, 000 X g for 5 minutes. The collection tube was emptied and the column assembly was recentrifuged for 1 minute with the microcentrifuge lid open (off) to allow evaporation of any residual ethanol. Minicolumn was carefully transferred to a clean 1.5 ml microcentrifuge tube and 50  $\mu$ l of nuclease free water was added to the Minicolumn. The mixture was incubated at room temperature for 1 minute and centrifuged at 16, 000 X g for 1 minute. The Minicolumn was discarded and DNA was stored at 4°C or -20°C.

## APPENDIX B

**Table B1:** DNA sequence variations in Vervet monkeys

Monkey number	Gene	Exon	Nucleotide change	A.A change	Type	polarity
311, 374,387, 389, 397,400, 402, 416,116, 236 & 215.	<i>HSF4</i>	5	c.1313 C>T	R116>R	Silent	Basic polar
400		10	c.1700 C>T	L245L	Silent	Neutral nonpolar
374,389, 387 & 398		14	c. 2227 C>T	P421>L	Missense	Neutral.nonpolar- Neutral.nonpolar
311,374,387, 416, 398 and 206	<i>CRYAA</i>	3	c.470 C>G	S134>W	Missense	Neutral polar-Neutral slightly polar
402,311& 387		3	c.567 G>C	K166>N	Missense	Basic polar-Neutral polar
374,394,389, 206,1077,215, 236, 400,402, 409, 371 and 416	<i>GCNT2</i>	1A	c.917 G>A	G212>G	Silent	Neutral.nonpolar
400,402,206 & 409			c.1049 T>C	H256>H	Silent	Basic polar
400,402, 206 & 409			c.1053 A>G	M258>V	Missense	Neutral.nonpolar- Neutral.nonpolar
416, 371, 409, 400,394,116, 1077,236&389			c. 1106 C>T	N275>N	Silent	Neutral polar
389, 400, 402	<i>GCNT2</i>	1B	c.754 G>A	V16>I	Missense	Neutral.nonpolar-

236 & 416						Neutral.nonpolar	
397, 398, 116, 215 & 409			c.1073 A>T	Y122>F	Missense	Neutral.polar-Neutral non-polar	
402 & 387	<i>GCNT2</i>	1C	c.275 T >C	S15>S	Silent	Neutral polar	
311			c. 301 G >A	S24>N	Missense	Neutral.Polar- .polar	Neutral
398, 371, 374			c. 344 T >C	S38>S	Silent	Neutral polar	
389, 371, 402, 374, 387			c.794 C > T	I188>I	Silent	Neutral nonpolar	
400, 409, 398, 371, 416, 397, 311			c.812 C >T	D194>D	Silent	Acidic polar	
	<i>GCNT2</i>	2	No mutations	No mutations	-	-	
374, 387, 397, 416 and 389	<i>GCNT2</i>	3	c. 1670 C/T	Y373>Y	Silent	Neutral polar	
	<i>LIM 2</i>	1	No mutations	No mutations	-	-	
		2	A/G	P66>P		Neutral polar	
		3	T/C	I118>T	-	Neutral nonpolar	-
			G/A	A127>T		Neutral nonpolar	-
		4	-	-	-	Neutral polar	
		5A	No mutations	No mutations	-	-	