DNA SEQUENCE VARIATION IN NORMAL PIGMENTATION

Premila Rozanne John

DNA SEQUENCE VARIATION IN NORMAL PIGMENTATION

Premila Rozanne John

A thesis submitted to the Faculty of Medicine, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Human Genetics.

Johannesburg, South Africa, 1999

ABSTRACT

Pigmentation is one of the most visible physical characteristics in humans and can be influenced by various factors, both environmental and genetic. The genes in this polygenic trait take part in the stimulation of and production of melanin. This study aimed to explore the role of the *MC1R*, *TYR* and *TYRP1* genes in normal pigment variation.

Variation in the *MC1R* gene was investigated in the Negroid and San populations, and a group of red-haired Caucasoid individuals. The two African groups had fewer non-synonymous than synonymous mutations. The F196L (7/59 chromosomes) and T314T (25/59) variants were significantly more common in the Negroids than the San ($p<0.05$ and $p<0.01$), who were relatively less diverse at the *MC1R* locus. The L50L variant (3/34) was significantly more common in the San ($p<0.05$). The Negroids and San did not share any alleles with the red-haired Caucasoids, who were all found to be either homozygous for one mutation or heterozygous for two different mutations. They revealed four novel mutations S83P, Y152X, A171N and P256P at low frequencies (1/14).

A random group of normally pigmented Negroid and Caucasoid individuals were investigated for sequence variation at the *TYR* and *TYRP1* loci. In the *TYR* gene S192Y ($p<0.05$) and R402Q ($p=0.01$) mutations, observed in 8/30 and 6/28 chromosomes, were significantly associated with the Caucasoids, while the Negroids did not vary at this locus. *TYR*, therefore, probably plays a significant role in pigmentation differences between Caucasoids and Negroids. The *TYRP1* gene in both populations, however, was found to lack significant variation.

In conclusion, this study indicates that variation at the *MC1R* and *TYR* loci may play a significant role in normal pigment variation in humans, but that the *TYRP1* gene appears to play a less significant role.

DECLARATION

I declare that this thesis is my own, unaided work. It is being submitted for the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university. I declare that the protocol has been cleared by the Committee for Research on Human Subjects (clearance certificate protocol numbers M970510 and M940711).

Premila Rozanne John

25th day of May, 1999

This work is dedicated to my mother and brother for all their support, encouragement and patience, and to my late father who taught me the value of good education and would have been proud of this achievement.

ACKNOWLEDGMENTS

I would like to first thank all the individuals who participated in this study by giving us blood samples, particularly the red-haired individuals. I am grateful to Dr. R. Sturm and Dr. B. Rana for making primer sequences for the *MC1R* study available to me prior to publication.

A big thank you to all my colleagues in the Molecular Genetics laboratory for all their help with the techniques, their support and advice. A special thank you to Dr. T. Lane and Ms. A. Turner for their help with the statistical analysis and to Ms. C. Padoa, Mrs. B. Morar and Ms. R. Kerr for their help with proof-reading and advice with this dissertation.

I would especially like to place on record my humble gratitude to my supervisor Prof. M. Ramsay for giving me the opportunity to do this project and for all her guidance throughout my research.

Finally, to the FRD, MRC and SAIMR, my heartfelt appreciation for the financial support that was provided for this project.

TABLE OF CONTENTS

LIST OF TABLES

4. DISCUSSION

LIST OF FIGURES

3. RESULTS

LIST OF ABBREVIATIONS

A	adenine
α-MSH	α-melanocyte stimulating hormone
APS	ammonium persulfate
ASP/ASIP	agouti signalling protein
AS	Angelman syndrome
ASO	allele specific oligonucleotide
BOCA	brown oculocutaneous albinism
bp	base pairs(s)
C	cytosine
°C	degrees Celsius
cAMP	cyclic AMP (adenosine 3',5'-cyclic monophosphate)
cDNA	complimentary DNA
CHS	Chediak-Higashi syndrome
Ci	Curie
cm	centimeter
dCTP	deoxycytidine-5'-triphosphate
DGGE	denaturing gradient gel electrophoresis
DHICA	dihydroxyindole-2-carboxylic acid
DIG	digoxygenin
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
ddNTP	dideoxyribonucleotide triphosphate
EDTA	ethylenediamininetetra-acetic acid
g	gram(s)
G	guanine
HPS	Hermansky-Pudlak syndrome

kb	kilobase(s)
l	litre
MC1R	melanocortin 1 receptor (or melanocyte stimulating hormone receptor)
mA	milliamps
ml	millilitre
mM	millimolar
mV	millivolts
M	molar
µl	microlitre
µg	microgram
µCi	microCurie
µmol	micromoles
µM	micromolar
mRNA	messenger RNA
MW	molecular weight
MSHR	melanocyte stimulating hormone receptor
ng	nanogram
nm	nanometer
OCA	oculocutaneous albinism
p	short arm of chromosome
P	P protein
PCR	polymerase chain reaction
%	percentage
pM	picomolar
pmol	picomoles
PWS	Prader-Willi syndrome
q	long arm of chromosome
RE	restriction enzyme
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid

rpm	revolutions per minute
ROCA	roufus oculocultaneous albinism
s	seconds
SDS	sodium dodecyl sulphate
ss	single stranded
SSC	sodium citrate solution
SSCP	single strand conformation polymorphism
T	thymine
TBE	tris borate
TE	tris EDTA
TEMED	N,N,N',N'-tetramethylethylene diamine
Tris	tris (hydroxymethyl) aminomethane
TYR	tyrosinase
TYRP1	tyrosinase related protein 1
TYRP2	tyrosinase related protein 2
U	units
UV	ultraviolet
V	volts
vs	versus
W	watts

1. INTRODUCTION

The earth is populated by a diverse group of human individuals who differ in many ways. One of the most visible forms of diversity is pigmentation of the skin, hair and eyes. Pigmentation has been widely studied for many years but not much is known about the detailed mechanisms involved in this process. Disorders of hypopigmentation have been investigated in an attempt to identify genes involved with pigmentation. Albinism is one of the most well studied hypopigmentation disorders and these studies have lead to the identification of several pigment genes. These loci are likely candidates that contribute to normal pigmentation variation.

The genetic complement and effects of the environment determine the phenotype of an individual. The pigmentation process is complex and can be studied at different levels. The first level involves the phenotype and changes in pigmentation related to various evolutionary theories. The next is the cellular level, in which the cells and tissues involved in pigmentation may be studied. Then, the biochemistry of the process of pigmentation, termed *melanogenesis*, which involves a complex pathway known as the *melanin biosynthetic pathway* that produces the pigment *melanin*. Finally, the molecular level, which involves the genes. In this study, the molecular basis of pigmentation will be investigated to add to the body of information that is being built up on normal variation in pigmentation.

1.1 Normal pigment variation

Disturbances of pigmentation in hypopigmentary disorders provide information that is necessary to decipher the process of normal pigment variation. Many shades of skin colour occur from the very dark Negroid populations to very light Caucasoid populations. Hair colour varies from black to red to blond, with various shades of colour in between. These differences may have evolved as an adaptation to the environments that humans are exposed to. The proposed biological basis of normal pigment variation, which has been studied extensively in the past decades, will be discussed in this section.

1.1.1 Melanogenesis

Skin, hair and eye pigmentation is the most visible variation in humans. This variation in pigmentation is due to the various ratios of the two different melanins (*phaeomelanin* and *eumelanin*) and the way in which these are packaged in the *melanosome*, the organelle in which melanogenesis occurs. Melanin is a polymer that is produced by *melanocytes*, the pigment producing cells of the body. This pigment has been proposed to have some important functions such as a camouflage, heat absorption, and protection of the skin from the dangerous rays of ultraviolet (UV) light (photoprotection) [Bolognia and Paweleck, 1988]. It has also been found to play a role in the scavenging of free radicals and in the routing of the optic nerve tracts [reviewed in Sturm *et al*, 1998]. The production of melanin occurs by a complex biosynthetic pathway, which involves many different components that have specific roles. The process of melanogenesis, including the structures of the skin, hair follicle and eye, and the biochemistry of the process will be discussed in the following sections.

1.1.1.1 Skin, hair and eye tissue structure

In order to understand melanogenesis it is important to understand the architecture of the skin, hair follicles and eyes and their pigmentation systems.

The skin is composed of two main layers, the *dermis* and the *epidermis* (Fig.1.1). The dermis is a connective tissue layer made of collagen, elastin and reticular fibres. It is supplied with blood, lymphatic vessels and sensory nerve endings. It also has sweat and sabaceous glands, hair follicles and hair. The epidermis is the thinner layer that is composed mainly of *keratinocytes*

(epidermal cells) and melanocytes. The keratinocytes are arranged in four layers: the *basal layer, stratum spinosum, stratum granulosum,* and *stratum corneum* (Fig.1.2). The keratinocytes are formed in the basal layer and move upwards from layer to layer as they develop and mature until they reach the stratum corneum, the top most layer. When they reach the stratum corneum they are dead keratinised cells, which are shed to balance the constant formation of new cells in the basal layer. The melanocytes make up approximately 10% of the cells in the basal layer [Robins, 1991]. The keratinocytes and melanocytes together give the skin its pigmented appearance.

The hair follicle is the epidermal tube from which the hair shaft grows (Fig.1.1). The base of the hair follicle forms a bulb in the dermis of the skin. Hairs are essentially tightly compacted and dead keratinised cells. The cuticle of the hair is the skeleton that anchors it firmly in position by fusing its cells with that of the follicle (Fig.1.3). The cortex, the bulk of the hair, consists of closely interlocking keratinised cells with air spaces in between. The medulla contains more loosely distributed keratinised cells with larger air spaces. Hair is generated during the hair follicles' active phase (*anagen*), and ceases during the quiescent phase, *telogen* [Robins, 1991]. During telogen the melanocytes, which are situated in the part of the hair-bulb closest to the dermal papilla, are also switched off [Robins, 1991; Ortonne and Prota, 1993]. Hair colour occurs as a result of the pigment situated in the keratinocytes of the cortex and medulla of the hair [Robins, 1991].

The eye is a complex organ composed of different layers (Fig.1.4) with unique functions. Some parts are pigmented. Eye colour is dependent on the amount of pigment deposited on the iris, in the stromal melanocytes of the iris and the posterior pigment epithelium. Brown eye colour is a result of large amounts of pigment, while less iris stroma pigmentation results in a blue, hazel or green appearance [Vaughan and Asbury, 1977; King *et al,* 1995]. The retinal pigmented epithelium (RPE) acquires melanin early during embryogenesis. Pigment is also present in the choroid below the retina. Pigment is necessary during the optic system development for the correct routing of the optic nerves to the lateral geniculate nucleus and for foveal development. The amount of pigment needed for normal routing is not known [Oetting *et al,* 1996].

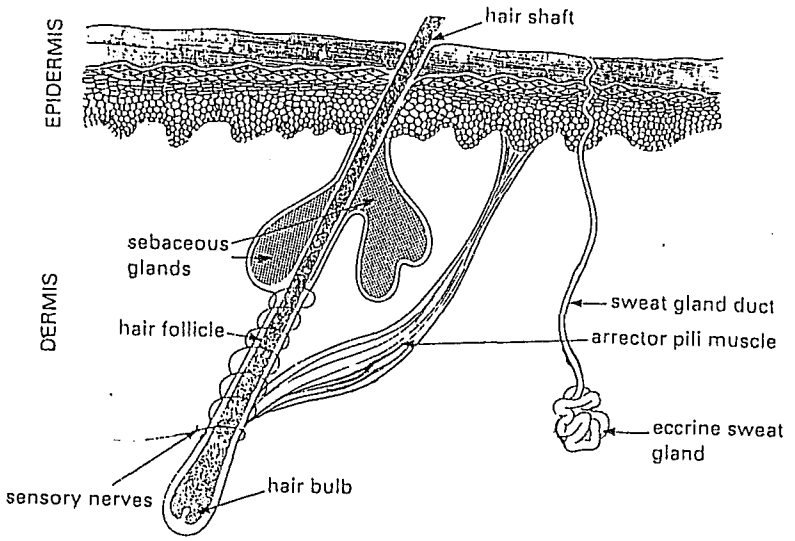


Fig.1.1: The architecture of the skin. This structure of the skin shows the position of the epidermis and the dermis, as well as the position of the hair follicle in the skin. [Figure modified from Robins, 1991].

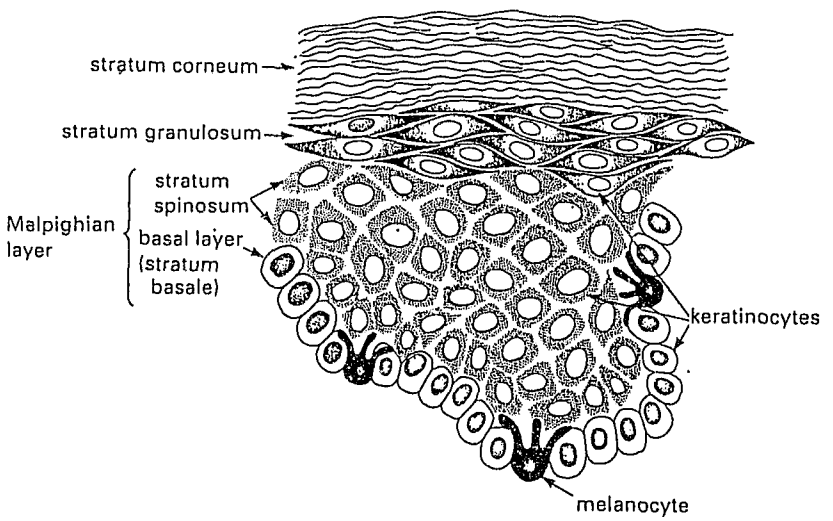


Fig.1.2: The four keratinocytes layers. A more detailed structure of the epidermis, showing the four layers of keratinocytes and the location of the melanocytes. [Figure from Robins, 1991].

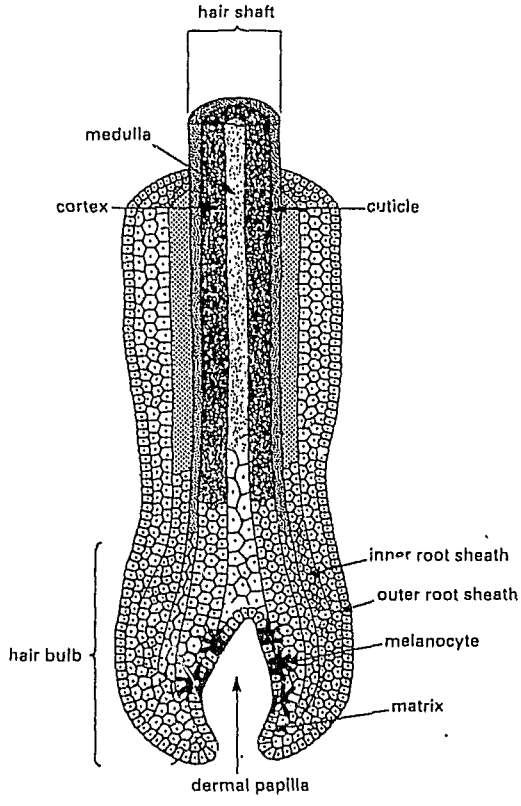


Fig.1.3: The structure of the hair follicle and hair. This figure shows the different keratinocyte layers in the hair and the position of the melanocytes in the hair-bulb. [Figure from Robins, 1991].

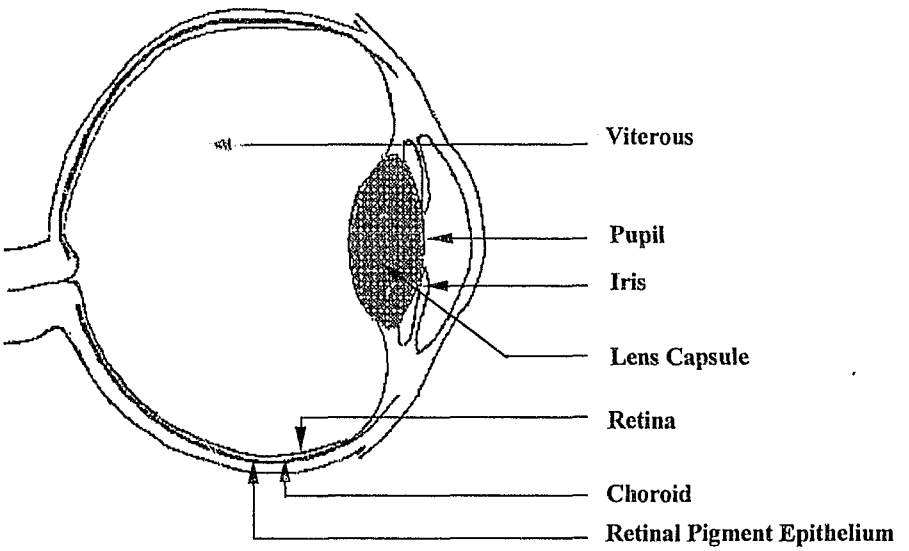


Fig. 1.4: The structure of the eye. This figure shows the positions of the parts of the eye involved in pigmentation. [Figure modified from Vaughan and Asbury, 1977].

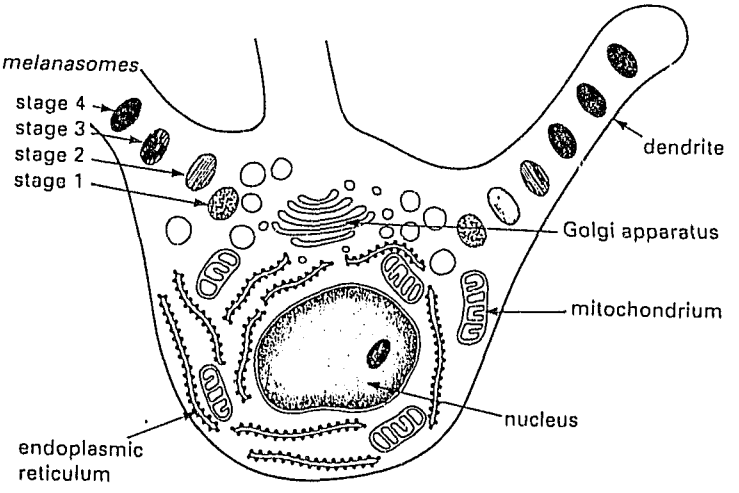


Fig.1.5: The dendritic melanocyte cell. This figure shows the structure of the melanocyte and illustrates the stages of melanosome development in the melanocyte. [Figure from Robins, 1991].

1.1.1.2 Melanocytes and melanosomes

Melanogenesis occurs in the melanocyte in membrane bound organelles, the melanosomes [Bolognia and Paweleck, 1988; Robins, 1991; King *et al*, 1995]. The melanocytes are dendritic cells (Fig.1.5) found in the epidermis of the skin and the hair-bulb, above the dermal papilla [Robins, 1991].

The origin of melanocytes

Melanocytes are mainly present in the basal layer of the epidermis and the matrix of the hair follicle. They are also found in the dermis, mucous membranes, the retina and the uveal tract of the eye, and the cochlea and vestibular labrynth of the ear [Bolognia and Paweleck, 1988]. Melanocytes in most areas, except those in the retina, are derived from precursor cells in the neural crest. The retinal melanocytes originate from the optic cup of the forebrain [Bolognia and Paweleck, 1988; Robins, 1991]. The iris and middle ear melanocytes also arise from the neural crest [Orlow, 1995]. In the embryo the neural plate gives rise to the neural tube and the band of cells of the neural crest. The neural crest gives rise to many cells including cells known as *melanoblasts*, which migrate to the head region along either side of the spinal chord to the skin. They then enter the dermis, epidermis and hair follicles where they differentiate to melanocytes. The dermal melanocytes decrease in number during gestation and virtually disappear at birth, after which melanin is primarily produced in the epidermis [Robins, 1991].

Melanosome production

The melanosome is produced from the smooth endoplasmic reticulum (SER) as a cytoplasmic membrane bound vesicle, and is referred to as the stage I premelanosome, which does not contain enzymes for melanin synthesis. This premelanosome then develops an internal matrix, but still contains no enzyme for melanin synthesis at which point it is a stage II premelanosome [King *et al*, 1995]. The stage II premelanosomes fuse with coated vesicles containing the enzyme tyrosinase. These vesicles are released form the Golgi apparatus in the cytoplasm of the melanocyte [Orlow *et al*, 1993; King *et al*, 1995]. Tyrosinase is activated by its glycosylation in the Golgi-associated endoplasmic reticulum (GREL) and the coated vesicle system [Pavel, 1993]. The association of the premelanosome with the coated vesicle leads to the synthesis of melanin in the premelanosome. This melanin is deposited uniformly within the matrix of the

melanosome, giving rise to the stage III premelanosome. Gradually, melanin fills the organelle until it is completely opaque. This is the mature stage IV melanosome. During this process the melanosome moves from the perinuclear region to the periphery of the melanocyte, where it is transferred to surrounding keratinocytes [King *et al*, 1995]. Fig.1.5 and Fig.1.6 show the four stages of the melanosome as it matures. The production of the melanosomes is the same in epidermal and follicular melanocytes. However, it must be noted that the melanosomes of the hair follicle are two to four times larger than the epidermal melansomes [Ortonne and Prota, 1993].

The melanins

Melanin in mammalian melanocytes is a heterogeneous bipolymer [Ando *et al*, 1995], which consists of the two types of melanin - the black/brown eumelanin and the red/yellow phaeomelanin. Both types of melanins are present in normal human epidermis [Thody *et al*, 1991]. Eumelanin and phaeomelanin differ in their chemical composition as well as the ultrastructure of their respective melanosomes [Sakai *et al*, 1997]. Eumelanin is produced by eumelanosomes and phaeomelanin is produced by phaeomelanosomes. The former are ellipsoidal and fibrillar while the latter are ovoid and particulate, as seen in the hair follicles of red-haired individuals [Pavel, 1993; Sakai *et al*, 1997]. Each melanosome has the capacity to produce both types of pigment and when they do, mixed melanins occur [Hearing and Tsukamoto, 1991; King *et al*, 1995]. The exact ratio of eumelanin to phaeomelanin in the mixed melanins varies and the mechanism that controls the determination of this ratio has not been found yet [King *et al*, 1995]. An eumelanin to phaeomelanin ratio of greater than 1 results in black/brown hair and that less than 1 results in yellow/red hair [Sakai *et al*, 1997]. At stage I of melanosome development both melanosome types are round in shape and have a membrane surrounding the microvesicle and filament. They both go through the maturation procedure as described previously. But, although eumelanin is deposited evenly in the eumelanosomes, phaeomelanin is deposited irregularly in the matrix of phaeomelanosomes [Witkop *et al*, 1983].

The epidermal melanin unit

The dendrites of the melanocytes in the skin are positioned so that each melanocyte is associated with between 20 and 40 keratinocytes [Robins, 1991]. This partnership has been termed the

epidermal melanin unit (Fig.1.7), which regulates pigmentation. The melanosomes are transferred to the keratinocytes by a dendritic process, in which the squamous cells phagocytise the melanosome laden dendritic tips of the melanocytes [Bolognia and Pawelek, 1988; Chakraborty and Pawelek, 1993]. In the hair follicles this association between melanocytes and keratinocytes is known as the *follicular melanin unit* [Ortonne and Prota, 1993]. The melanosomes entering the keratinocytes can be arranged in two configurations: either as single particles (non-aggregated) or as a package of a few melanosomes within a membrane bound vesicle (aggregated form), as shown in Fig.1.7. According to studies, the small ellipsoidal melanosomes aggregate while the large ones do not [Witkop *et al,* 1983; Robins, 1991]. Within the keratinocytes the aggregated melanosomes undergo degradation by hydrolytic enzymes contained in them, which degrade the protein and lipid components of the melanosome but not the melanin. The latter occurs as the basal keratinocytes are pushed upwards in the epidermis. The non-aggregated forms seem more resistant to degradation and remain relatively intact through the process [Robins, 1991]. This process occurs in the follicle as in the epidermis [Ortonne and Prota, 1993]. Hence, the pigmented appearance of the skin and hair. The melanocytes in the eye, however, retain the melanosomes [Boissy, 1988 cited in Sturm *et al,* 1998].

1.1.1.3 The biochemistry of melanogenesis

The process of melanin biosynthesis is a stepwise procedure of various reactions that take place in a synchronised fashion in the melanosome of a normal individual to produce melanin. Various studies have been carried out to reveal the components of this pathway and the role that each plays in the melanocyte. Our present knowledge of the basis of pigment production is summarised in Fig.1.8 and will be discussed in detail in this section.

Stimulation of melanogenesis

Melanogenesis is triggered when an external stimulus is received by the melanocyte. This external stimulus is *α-melanocyte stimulating hormone* (α-MSH), a tridecapeptide, that is derived from the protein *proopiomelanocortin* (POMC) [reviewed in Eberle, 1988, cited in Robbins *et al,* 1993]. α-MSH is released from the pituitary into the blood and is also produced peripherally in the skin by keratinocytes [Barsh, 1996; Valverde *et al,* 1996]. Its main function is

the stimulation of pigment production in melanocytes and melanoma cells and it has been found to affect melanocyte proliferation [Eberle, 1980]. α-MSH binds to receptors on the surface of the melanocytes known as the *melanocortin 1 receptors* (MC1R) or the *melanocyte stimulating hormone receptors* (MSH-R), which are present in large quantities on the surface of the melanocytes. MC1R is a seven transmembrane G-protein coupled receptor [Grantz *et al*, 1994; Frandberg *et al*, 1998] which, on binding with α-MSH, transduces the signal via GTP-binding protein to adenyl cyclase. This results in an increase in the production of cAMP, which is thought to activate the enzyme *tyrosinase*, involved in the rate limiting step of melanogenesis [Ortonne and Prota, 1993; Frandberg *et al*, 1998]. The intracellular loops of G-protein coupled receptors are believed to be involved in the coupling of GTP-binding proteins [Frandberg *et al*, 1998]. Studies have shown that in response to the hormone-receptor binding, the level of tyrosinase mRNA and its activity increase [Jackson, 1993a; Rungta *et al*, 1996]. Rungta *et al* (1996) showed that transcription is affected, as tyrosinase mRNA increased at a slow rate in mouse melanoma cells in response to α-MSH.

Tyrosinase and the start of melanin biosynthesis

Tyrosinase is a copper containing enzyme that is essential for melanogenesis as it catalyses the rate limiting reaction of the melanin biosynthetic pathway. Tyrosinase also takes part in two reactions further along in the pathway [Koer and Pawelek, 1982; Tripathi *et al*, 1992a]. The rate limiting step is the hydroxylation of the amino acid tyrosine to *L-3,4-dihydrophenylalanine* (DOPA) as shown in Fig.1.8. Tyrosine is transported across the melanocyte membrane by a Na^+-independent transport system, the L (leucine) system [Pavel, 1993]. Tyrosine uptake and melanin production are not fully correlated as hair follicle melanocytes that do not produce pigment take up normal amounts of tyrosine. Tyrosine uptake was also found to be greatest in red-haired individuals, which shows that it is not directly correlated with eumelanin synthesis in humans [Pavel, 1993]. The first two steps of melanogenesis are the same for the synthesis of both types of melanin, with tyrosine being the first component. The DOPA that is formed in the first step is oxidised to *DOPAquinone* in the presence of tyrosinase, or at a much slower rate in its absence [King *et al*, 1995]. At this point the pathway can lead to the production of either eumelanin or phaeomelanin.

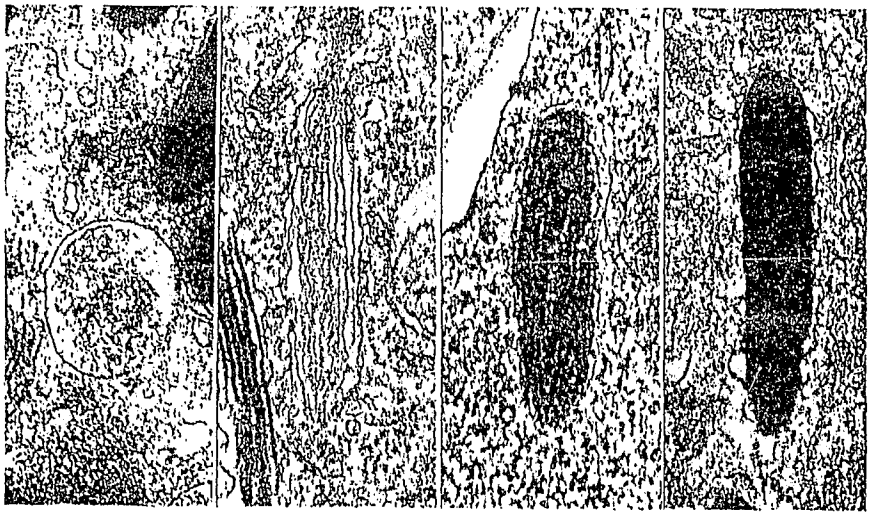


Fig.1.6: Stages of melanosome development. The melanosome develops from stage I premelanosome, containing no melanin, to a stage IV melanosome, which is completely full with melanin. [Figure from Witkop *et al*, 1983].

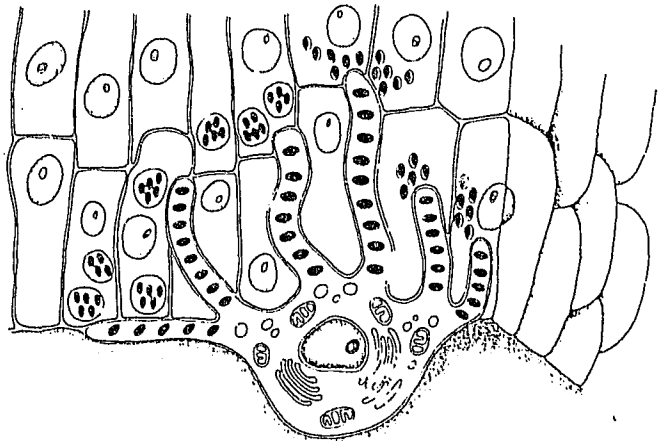


Fig.1.7: The epidermal melanin unit. This structure consists of the melanocyte that is surrounded by keratinocytes. This figure also shows the two ways in which the melanosomes are transferred to keratinocytes, as the aggregated form (left) and the non-aggregated form (right). [Figure from Robins, 1991].

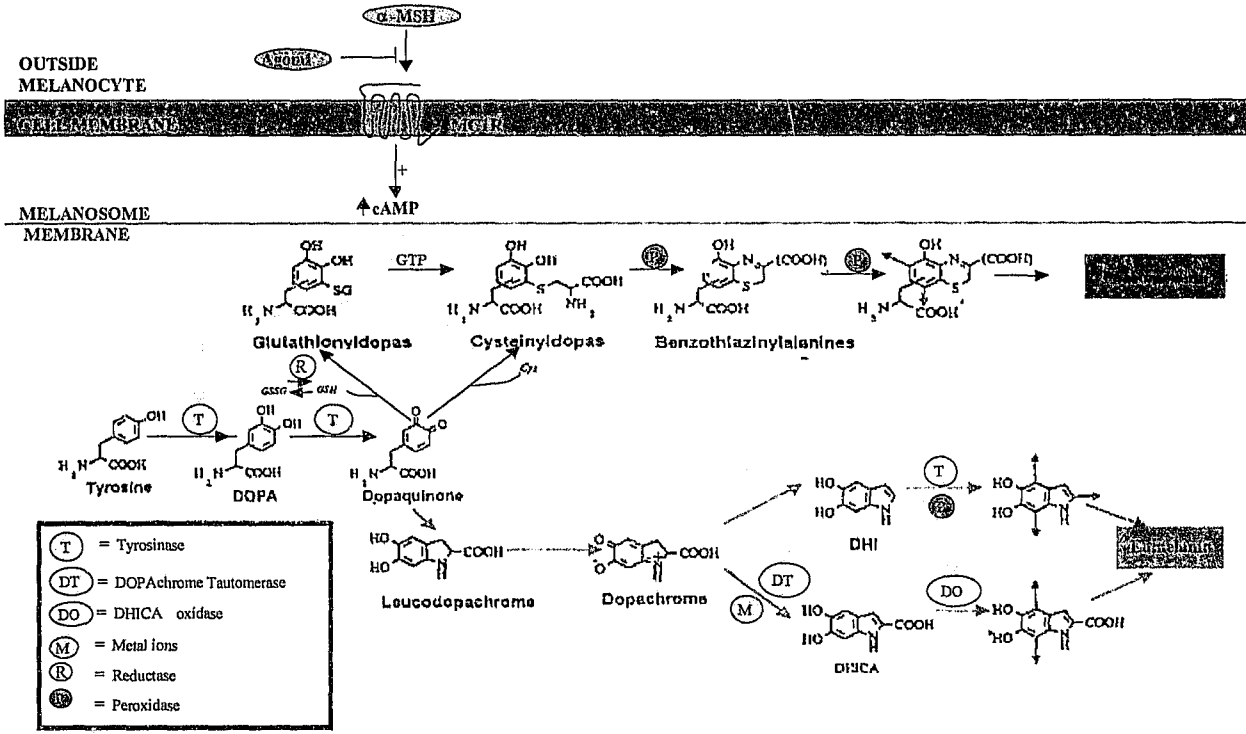


Fig.1.8: The melanin biosynthetic pathway. The diagram shows the factors necessary for melanin synthesis. It shows the stimulatory effect the binding of α-MSH to MC1R has on tyrosinase, DHICA oxidase and DOPAchrome tautomerase (dotted arrows). The red arrows show the pathway to phaeomelanogenesis and the brown arrows to eumelanogenesis. A key for each enzyme or other component necessary for the individual reaction is presented in left side of the diagram. [King *et al*, 1995; Barsh, 1996; Prota, 1997].

Eumelanogenesis

Leukodopachrome is formed from DOPAquinone, which is further processed to give *DOPAchrome*. At this point there are two possible routes that can lead to the production of eumelanin. The first one involves the non-catalytic decarboxylation of DOPAchrome to *5,6-dihydroxyindole* (DHI), which can form *indole-5,6-quinone* in the presence or absence of tyrosinase or peroxidase. This then forms eumelanin. The second path to eumelanogenesis involves two additional enzymes. The first is *DOPAchrome tautomerase* (DT) which catalyses the reaction for the formation of *dihydroxyindole-2-carboxylic acid* (DHICA) in the presence or the absence of certain divalent metal cations such as cobalt, magnesium and iron ions [Prota, 1993; King *et al*, 1995]. In the absence of these metal ions, however, the former route is favoured and DHI is mainly formed with only about 5% DHICA [Prota, 1993]. DHICA is then processed to *indole-5,6-quinone carboxylic acid* in the presence of another enzyme, *DHICA oxidase*, which is then processed to eumelanin [King *et al*, 1995]. The structure and functions of the eumelanin derived from DHI and from DHICA may be different, as the colour of the eumelanins is different. The eumelanin derived from DHI is said to be black while that from DHICA is yellowish brown [Hearing and Tsukamoto, 1991].

Phaeomelanogenesis

In the phaeomelanogenic pathway DOPAquinone, in the presence of sulfhydryl groups such as glutathione or cystein, gets converted to *cysteinylDOPA* and through a number of additional steps phaeomelanin is produced (Fig. 1.8) [King *et al*, 1995; Prota, 1997]. The availability of the sulfhydryl groups is the trigger that determines the production of phaeomelanin [Prota, 1993].

α-MSH and eumelanogenesis

Studies have shown that α-MSH promotes the production of eumelanin. The binding of the hormone to MC1R leads to the increased production of cAMP, which increases the production of tyrosinase to levels that favour eumelanogenesis [Frandberg *et al*, 1998]. This has also been shown to be the case with mouse melanoma cells [Hoganson *et al*, 1989]. Lower tyrosinase concentrations have been found to favour phaeomelanogenesis [Tobin *et al*, 1994; Frandberg *et al*, 1998]. α-MSH increases the expression of TYR (tyrosinase), TYRP1 (DHICA oxidase) and TYRP2 (DOPAchrome tautomerase) proteins post-transcriptionally [Abdel-Malek *et al*, 1995].

The agouti protein and phaeomlenogenesis

The *agouti signalling protein* (ASP or ASIP) is an antagonist to α-MSH for MC1R binding [Jackson, 1993a], in both human and mouse, resulting in the inhibition of tyrosinase activity [Suzuki *et al,* 1997]. It is produced by the dermal papillae cells in the hair follicle [Sakai *et al,* 1997] and can be described as a paracrine signalling molecule that determines which pigment the hair follicles produce [Wilson *et al,* 1995]. Wild-type mouse hair contains both pigment types, eumelanin and phaeomelanin. The tips and bases of the hair are black with a transverse band of yellow phaeomelanin in between. This is caused by a change from eumelanogenesis to phaeomelanogenesis midway during the growth of hair [Jackson, 1993a]. The model that has been presented shows that when the agouti protein binds to MC1R it prevents α-MSH from binding, which is less favourable for eumelanogenesis. Hence, there is a production of phaeomelanin, which causes the hair of the mice to turn yellow. This switch from eumelanin synthesis to phaeomelanin synthesis is affected if mutations occur in the genes that produce the agouti protein or the MC1R protein [Jackson, 1993b]. Studies on mice have shown that mutations in the *a* gene can produce dominant alleles which result in the continuous production of the agouti protein and hence the binding of it to the MC1R and a continuous production of phaeomelanin [Jackson, 1993b]. The latter results in the yellow mouse coat colour phenotype.

The mechanism of the switch from eumelanogenesis to phaeomelanogenesis is not completely known. Furumura *et al* (1998a) investigated the effects of ASP on murine melan-a (albino) cells and found the upregulation of certain genes: a DNA replication control protein, a basic helix-loop-helix transcription factor, and a novel gene that is expressed in retinal cells (function unknown). There is a down regulation of the *TYR* and *TYRP2* genes, known to encode the tyrosinase and DOPAchrome tautomerase enzymes. Hence, the switch may be more complex than has been thought. This mechanism may also occur in humans when phaeomelanin is produced and the ASP in humans may have the same effects as in the mouse [Furumura *et al,* 1998a].

Tyrosinase activity in different hair colours

A number of studies have been carried out to determine the different activities of this enzyme in different types of hair. Tyrosinase activity can be assayed, according to King and Witkop (1976),

by determining the rate of formation of [^{3}H]OH with the oxidation of L-3,5-^{3}H-tyrosine to L-5-^{3}H-DOPA, which measures the rate of tyrosine hydroxylase activity of tyrosinase. Tyrosinase activity was found to vary considerably in different groups of individuals with different hair colours [King and Witkop, 1976]. The activity was found to increase in hair-bulbs from light to dark hair, as would be expected. However, red-haired individuals had the highest tyrosinase activity [King and Witkop, 1976; King *et al*, 1978; Lloyd *et al*, 1987]. In many studies on epidermal melanocytes, high tyrosinase activity has been associated with eumelanogenesis and phaeomelanogenesis has been associated with low tyrosinase activity. Therefore, there may be a different mechanism in follicular and epidermal melanocytes. The concentration of eumelanin and phaeomelanin in the hair is much greater than in the epidermis, although this may be due to higher water content in the skin [Thody *et al*, 1991]. In mice, low tyrosinase activity is associated with phaeomelanogenesis, suggesting a different mechanism in humans. Thus, there is still much about the process of phaeomelanogenesis and the switch that controls the process that is not completely understood.

1.1.2 Phenotypic variation

The diversity of pigmentation of the different populations in the world (Fig.1.9) is due to genetic differences influencing the basic biochemical and physiological pathways involved in pigmentation as well as modification by environmental factors. UVR exposure is one of the biggest determinants of the variation in skin colour. Some differences have occurred as a result of adaptation of humans to different environmental factors as they moved and settled in different areas of the globe. The biochemical and physiological basis of these differences will be discussed.

1.1.2.1 Skin reflectance test

Skin colour can be assessed with the use of a skin reflectance measurement. This is a method whereby the light reflected by each skin sample is measured and used as a reliable and quantitative measurement of skin colour. The procedure involves an EEFL reflectance spectrophotometer, which consists of a galvanometer and a detachable reflectometer head. Light from a lamp in the reflectometer head directs light through a chosen filter and then through an

aperture onto the skin surface. Thus, different wavelengths of light can be tested. Some of the light is absorbed and some is reflected. The reflected light is picked up by a photocell and measured on the galvanometer unit. Lighter skin will absorb less light and reflect more and darker skin would absorb more light and reflect less, and thus a gradient of reflected light is constructed. The preferred site of skin reflectance measurements is the medial aspect of the left upper arm as it is easily accessible and is unlikely to have had much UV exposure, and therefore less likely to be tanned [Kromberg, 1986]. The reflectance values are expressed as percentages of the light reflected. Other suitable methods of skin colour evaluation or for quantitating skin colour are reflectance colorimetry and image analysing using sensitive colour video cameras [Pierard, 1998].

1.1.2.2 Sun reactive skin typing

In 1975, the concept of sun reactive skin typing was developed to be able to classify individuals with white skin to select the correct initial doses of UV-A for the treatment of psoriasis [Fitzpatrick, 1988]. This is a skin disorder that usually appears as inflamed, swollen skin lesions covered with a white silvery scale [National Psoriasis Foundation, 1998, internet]. Later, brown and black individuals were included in the classification [Fitzpatrick, 1988]. Table 1.1 lists the basic skin types classified on the basis of their sunburn and tanning ability [Ortonne, 1990; Abdel-Malek *et al*, 1994]. Sunburnt cells are keratinocytes that have undergone apoptosis as a result of UV irradiation. The long-term effects of UV exposure are photoaging and photocarcinogenesis. UV irradiation causes the melanocytes to arrest at the G2 phase of the cell cycle. This was also seen in fibroblasts and squamous carcinoma and melanoma cells. The arresting may allow for DNA repair and to stimulate melanogenesis to protect the melanocytes from further DNA damage. The G2 phase is one in which tyrosinase is predominantly expressed [Abdel-Malek *et al*, 1994]. Following UV-B exposure the melanocytes of different skin types have been found to be larger, more dendritic and to be greater in number [Tobin *et al*, 1994]. They found that there was no difference in the melanocyte number and distribution between skin types I/II (light skin) and III/IV (darker skin) following UV-B exposure.

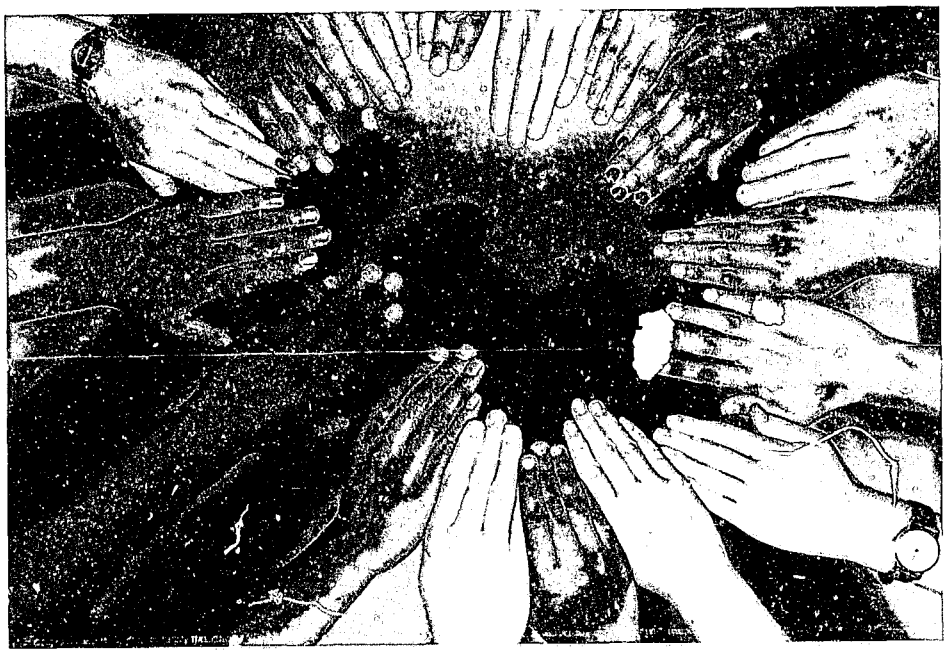


Fig.1.9: Diversity in skin pigmentation. This photograph shows the diversity in skin pigmentation in various individuals. This diversity has been thought to have occurred due to evolutionary forces.

Table 1.1: The skin reactive types [Fitzpatrick, 1988]

Unexposed skin colour	Skin type	Sunburn	Tan
White	I	Yes	No
	II	Yes	Minimal
	III	Yes	Yes
	IV	No	Yes
Brown	V	No	Yes
Black	VI	No	Yes

The concentration of the two melanins may vary in the different skin types. Thody *et al* (1991) demonstrated that phaeomelanin concentration varied but showed little relationship to skin type I, II and III, except that the only individual with red hair in the study had significantly high levels of phaeomelanin. However, eumelanin does have an association to skin type, as its concentration was lowest in skin type I and higher in types II and III. The relationship was the same with tyrosinase synthesis and skin type. An increase in both melanins was observed in all skin types when irradiated with UV-A, which was not related to skin type when considering phaeomelanogenesis. The greatest increase in eumelanin was seen in types II and III [Thody *et al*, 1991].

1.1.2.3 The biochemical and physiological basis of skin and hair colour diversity

The various colours of skin, hair and eyes are due to differences in the physiological and biochemical make up of these tissues in different people. The total number, size and packaging of melanosomes in the epidermal melanin unit determine skin colour. The rate of melanosome formation and the rate of melanosome transfer in keratinocytes [Robins, 1991], as well as the ratio of eumelanin to phaeomelanin, play a role.

Previous studies have shown that the critical factor for skin colour variation is the quality and quantity of melanosomes in the epidermis of the skin. In the melanocytes the Caucasoids have few melanosomes, with hardly any in stages III and IV. Mongoloid individuals have numerous stage II, III, and IV melanosomes while the Negroids have only stage IV melanosomes [Robins, 1991]. Mongoloids and Negroids have between five and eight to ten times more melanosomes respectively than Caucasoid individuals [Pavel, 1993]. In the keratinocytes the melanosomes can occur as membrane bound aggregates or be singly dispersed. The Caucasoids and Mongoloids have the aggregated melanosomes (Fig.1.10). In Mongoloids melanosome complexes are more tightly packed. Negroid individuals, on the other hand, have larger melanosomes that are more abundant in the basal keratinocytes [Robins, 1991]. Australian aboriginal people have a similar pattern to Negroids. The skin of Negroid babies, who have hardly had any exposure to UV light, contained more melanosomes that were more heavily melanised than other groups, proving that UV light was not the cause of these differences.

Hair colour varies widely from blond to black. These differences are brought about by differences in the melanosomes of the hair. Melanosomes were studied in the hair cortex; the inner, middle and outer areas were examined in different individuals. The Europeans had very low densities of melanosomes in all three areas, while Chinese had high quantities of melanosomes in the inner and intermediate areas and the Indians had the most in the outer areas. Surprisingly, Negroid individuals had lower levels in the inner and outer regions [Robins, 1991], unlike the epidermal melanocytes. These findings support different mechanisms in the hair follicle and epidermal melanocytes. In blond hair there are smaller and fewer melanin granules that are not fully melanised when compared with dark hair. In individuals with red hair the melanocytes usually contain mainly phaeomelanosomes which are spherical and contain spotty and granular melanin deposits, while other red-haired individuals have both eumelanosomes and phaeomelanosomes. Individuals with black and brown hair have eumelanosomes at various stages of melanisation and their ultrastructure is the same as that seen in the epidermis of Caucasoid and Negroid individuals. The melanosomes in the keratinocytes are singly distributed, as in dark skin, but in brown hair the melanosomes are smaller [Ortonne and Prota, 1993].

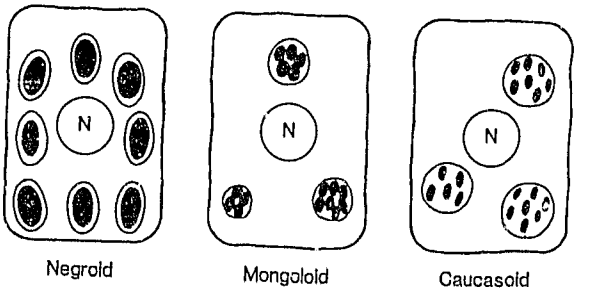


Fig.1.10: Melanosome structure. The melanosomes in the Negroid are large and singly dispersed. The lighter pigmented individuals have smaller melanosomes that are aggregated. In Caucasoids the melanosomes are more sparsely distributed aggregates. [Figure from Robins, 1991].

1.1.2.4 Age and melanogenesis

The ageing process in all individuals eventually results in white or grey hair and darker skin because of UV exposure. This process of ageing has been studied, and the changes in pigmentation seen with increased age have a physiological and biochemical basis.

Surprisingly, tyrosinase hair-bulb activity . . been found to increase with age, with a trend towards maximum activity at middle age. An early study by King and Witkop (1976) had shown that grey-white hair bulbs did not have any tyrosinase activity which was also seen in white hair-bulbs from individuals studied later [Lloyd *et al*, 1987]. Therefore, the tyrosinase activity increases to maximum at middle age and decreases from that point. Melanocytes in senile grey hair-bulbs have been found to be normal or reduced [Lloyd *et al*, 1987; Robins, 1991; Ortonne and Prota, 1993] and have very few melanosomes and little melanogenic activity. A decline in melanocyte number has also been observed in the epidermis; however, although the hair looks grey the skin hardly looses any colour, which may be due to a greater functional activity of older skin melanocytes and the accumulative sun exposure of them through the years [Robins, 1991]. α-MSH binding sites, normally present in pigmented hair follicles, have also been found to be absent in senile white hair follicles [Ortonne and Prota, 1993]. Greying has been said to appear earlier in dark- than in light-haired individuals, but fairer rather than dark-haired subjects become more completely grey. This contradiction can be explained by the fact that first signs of greying are more noticeable against a darker background. However, greying is less common in Negroid individuals. This process is usually irreversible, but a return to pigment has been known to occur [Ortonne and Prota, 1993].

1.1.3 Evolution of pigment variation

The variation of pigmentation in different geographical areas is thought to be a result of adaptation to climatic and other environmental changes in different parts of the world. The *vitamin D hypothesis* has been put across to explain the differences in skin colour across the earth as a result of selection. The climatic differences, in particular ultraviolet radiation (UVR), play a role in skin cancer. The adaptation of the human body to these climatic changes may have

been to prevent skin cancer in sunny areas and maintain vitamin D levels in the body in regions with less sunshine hours.

1.1.3.1 The vitamin D hypothesis

A paper by Loomis (1967) suggested a correlation between human pigment variation, latitude and vitamin D synthesis. Vitamin D refers to a group of chemically similar sterols that prevent and alleviate rickets, a disease of bone formation (where a disturbance occurs at the endochondral centres of ossification) in children and growing animals. There are two types that are of importance: ergocalciferol (vitamin D_2) and cholecalciferol (vitamin D_3) [Fraser, 1980]. In 1922, Hess discovered that sunlight could cure rickets [Loomis, 1967]. Vitamin D_3 is naturally occurring, and is formed by the action of sunlight on *7-dehydrocholesterol* in skin. The provitamin 7-dehydrocholesterol is present at a concentration of $1\mu g/cm^2$ in the epidermis [Fraser, 1980]. It absorbs UVR to produce the previtamin D [Avioli, 1979; Fraser, 1980], that is then slowly converted to vitamin D_3 without UV activation [Ichii-Jones *et al*, 1998]. This vitamin D_3 is absorbed into the subepidermal microcirculation and transported in the plasma, bound to a globulin, to the liver and other organs [Avioli, 1979].

The shortest wavelength of the sun's UV that reaches the earth is 290nm, but only the bottom 20nm to 30nm of the radiation that reaches the earth's surface plays a role in vitamin D_3 synthesis. Only UVR less than 310nm can be absorbed by the skin [Neer, 1975]. The intensity of the irradiation differs with latitude and season. At 53°N the intensity of light is sometimes 100 times more in summer than winter. The shorter wavelengths have a lower penetrating power and are reduced in pigmented skin. Melanin acts as a barrier against UV light. In Caucasoids 3% of the 290nm light penetrates the epidermis whereas 44% of the 320nm light penetrates, while in Negroids none of the shorter wavelength is able to penetrate and only 3% of the 320nm can [Fraser, 1980]. Recent research has shown that UV radiation is higher in the Southern Hemisphere than the Northern Hemisphere at a similar latitude [Relethford, 1997].

According to Loomis (1967) the vitamin D levels in the body must be maintained at a certain level (0.01 to 2.5mg/day) in order to avoid rickets. Thus, the fact that lighter skinned individuals are more prevalent in the northern latitudes is to enable them to absorb more UVR to maintain

optimum vitamin D levels. Individuals closer to the equator have darker skin, do not absorb too much UVR and have some protection against skin cancer [Loomis, 1967]. Skin reflectance is lowest at the equator (darker skin) [Relet..ford, 1997]. Too much UVR could lead to excessive vitamin D production, which is toxic [Loomis, 1967]. This may be the reason that skin tans during summer, to prevent the over production of vitamin D. A study by Relethford (1997) showed that human skin colour is darker in both men and women in the Southern Hemisphere when compared to the Northern Hemisphere at equivalent latitude. This is thought to be an adaptation of the skin to the different levels of UVR to produce optimum levels of vitamin D. As the populations moved from Africa towards the North they had to adapt to the climate to be able to produce more vitamin D to prevent rickets. Another example of adaptation to the environment is the darkening of skin as babies grow older, as more vitamin D is needed in the earlier years of life for bone development [Loomis, 1967]. Fig.1.11 shows the geographical distribution of skin colour variation in the different areas of the earth [Cavalli-Sforza *et al*, 1994].

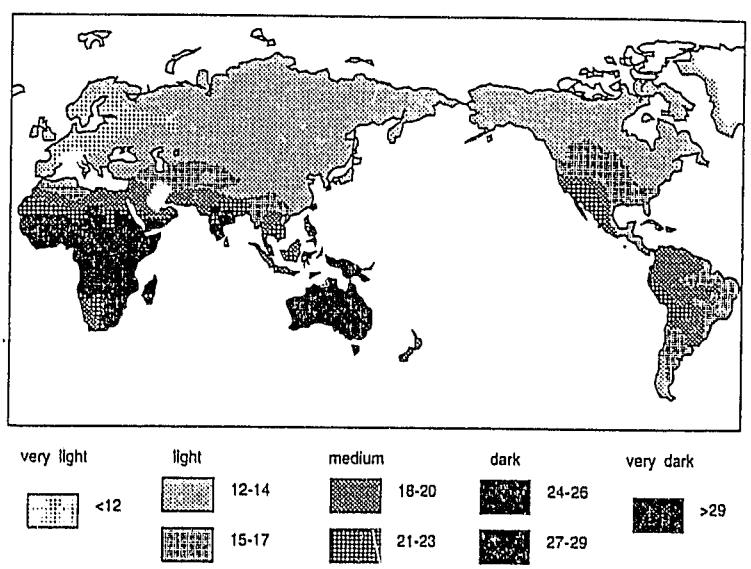


Fig.1.11: The geographical distribution of the pigmentation variation. The lighter pigmented individuals live in the northern areas of the earth, w'.ile darker individua'.s live towards the so.:th. [Figure from Cavalli-Sforza *et al*, 1994].

Studies have been carried out that do support this hypothesis. In recent years there has been a major migration of populations to the northern areas of the globe from the south and many darkly pigmented individuals live in the northern latitudes at present. A study in Philadelphia found a common occurrence of vitamin D deficiency in black mothers and infants [Bachrach *et al*, 1979], while another study by Awumey *et al* (1998) has shown that Asian Indians residing in the United States have a high risk of vitamin D deficiency and rickets. These studies support the vitamin D hypothesis. However, despite the evidence supporting the hypothesis, other authors have suggested that UV light is not the only factor that accounts for the geographical pattern of skin colour distribution [Fraser, 1980] and that "racial pigmentation" does not interfere with the production of vitamin D [reviewed in Hill, 1992]. It is also unlikely that lighter skinned people would synthesise toxic levels of vitamin D when living in tropical climates [reviewed in Jones *et al*, 1994].

1.1.3.2 Skin Carcinogenesis

Skin cancer is one of the most common cancers in the world today. There are three types of skin cancer: basal cell carcinoma, squamous cell carcinoma and malignant melanoma [Skincancer facts, 1998, internet]. Basal cell skin cancer affects the basal layer of the skin and squamous cell carcinoma affects the epithelial layer of the skin, including the keratinocytes. Malignant melanoma, the most dangerous of the three types [Skincancer facts, 1998, internet], is defined as a tumour derived from melanocytes in the epidermis of the skin [Crombie, 1979; Rippey and Rippey, 1984]. Skin cancer has a higher frequency in lighter pigmented skin when exposed to UVR. These and other findings will be discussed here, in relation to different skin and hair phenotypes.

Melanin and photoprotection

The main function of melanin is photoprotection, the protection of the skin from damaging effects of UVR. Melanin acts as a scavenger of reactive oxygen intermediates, which cause single strand breaks and base damage in DNA [Abdel-Malek *et al*, 1994]. Of the types of melanin, eumelanin is photoprotective, while phaeomelanin generates free radicals in response to UVR, which can damage the skin [Valverde *et al*, 1995]. The intermediates of eumelanogenesis, DHI and DHICA, also play a protective role in the skin. They have been found to protect the skin

by scavenging oxygen species, inhibiting lipid peroxidation, as well as protecting the skin from UV damage [reviewed in Sturm *et al*, 1998].

Population studies

The incidence of melanoma in the different parts of the world varies according to latitude and skin type. The occurrence of melanoma is higher in lighter skinned individuals than in darkly pigmented individuals. Malignant melanomas were found to be 20 times more frequent in American whites than American blacks [Reintagen *et al*, 1982]. Crombie (1979) noted that white individuals had a wide range of melanoma incidences and that individuals that are not white had much lower incidences. He looked at individuals from different latitudes in the world. In a study carried out in South Africa, the Rippeys (1984) reported that melanoma occurs from 2.5 to 6 times more frequently in whites than in blacks, but the lowest incidences were seen in the coloured population.

Different studies on white populations in South Africa, Australia and Norway show that the incidence of malignant melanomas is highest near the equator. This is consistent with the theory that malignant melanomas are due to UVR exposure to pigment-unprotected skin. Hence, the incidence of melanoma in whites varies with latitude, the occurrence being greater near the equator [Rippey and Rippey, 1984].

The types of lesions also differ between the whites and non-whites. The whites were found to have smaller lesions, less than 20mm in diameter situated on any part of the body. The blacks had lesions as large as 100mm in diameter, mainly on the soles of the feet and the palms of the hands, which are not necessarily exposed to the sun [Crombie, 1979; Reintgen, 1982; Rippey and Rippey, 1984]. Hence, the lesions in the Negroid individuals are most likely not due to UVR exposure, but probably another factor. It was reported that these tumours are most frequent on the weight-bearing areas of the sole, indicating that trauma is the most important factor [Crombie, 1979; Reintgen, 1982]. A small percentage of these tumours were found to be due to pre-existing moles [Reintgen, 1982]. Therefore, the incidence of melanoma in Negroid individuals can not be related to latitude, unlike the white individuals.

Individuals with albinism have a high risk of skin cancer because of their decreased epidermal melanin content. Studies have shown that South African subjects with albinism have a thousand times the incidence of skin cancer than normal Negroid individuals and ten times that of American Caucasoid individuals [Robins, 1991]. The rate of skin cancer in albinism is not uniform in Africa. In Nigeria, Okoro (1975) found that all the individuals with OCA studied, over the age of 20 years, had malignancies or premalignant lesions. Rose (1973), on the other hand, reported that skin cancer was not frequent in the Transkei in South Africa, where only 1 case was found in the 458 individuals with albinism that were studied [cited in Kromberg *et al*, 1989]. A study by Kromberg *et al* (1989) on albinism in Johannesburg revealed an overall prevalence of 23.4% skin cancer cases in 111 Negroid individuals with albinism. This frequency increased with age; from 6.0% in individuals below 10 years of age (2 of 32 individuals) up to 100% in subjects over 50 years of age (5 of 5 individuals). The head was the most commonly effected region of the body. No melanomas were observed in the individuals with albinism who were studied. Sun exposure resulted in basal and squamous cell carcinomas in these individuals [Okoro, 1975; Kromberg *et al*, 1989]. The prevalence of skin cancer in individuals with albinism in Africa, as a whole, is inconsistent, although the incidence is greater near the equator where the intensity if UVR is the greatest. For example, in South Africa, skin cancers were increased in Sotho-Tswana individuals with albinism, who live nearer to the equator, than in the Nguni [Kromberg, 1987].

1.1.3:3 Albinism in Africa

Albinism is common in Africa, where it occurs approximately two to three times more frequently in South African Negroids than in European Caucasoids [Roberts *et al*, 1986]. OCA2 is the most common type of OCA in the world because it has a high frequency in equatorial Africa. The phenotype of OCA2 is distinctive in African Americans and in African individuals [King *et al*, 1995].

Heterozygous advantage may occur with albinism in Africa, though the selective agent remains unknown. A classic example of heterozygous advantage is the HbS allele for sickle cell anaemia and malaria. The HbS allele has been found in high frequencies in certain parts of Africa even though individuals homozygous for HbS have the sickle cell anaemia disorder. This is so because individuals heterozygous for this allele are protected against infection by the malarial

parasite [Tobias, 1974]. Hence, in certain parts of Africa that have a high incidence of malaria the HbS heterozygotes have a survival advantage and, therefore, the frequency of this allele in those African populations is increased.

Likewise, there could be a selective advantage to carrying an allele causing a type of albinism, even though homozygotes have a disadvantage. Certain African individuals with one mutant allele may have lighter pigmentation than other random individuals from the same population, which in some communities is considered more attractive, and they may be selected for marriage more often. Hence a social heterozygote advantage [Roberts *et al*, 1986], where the lighter pigmented individuals parent more children, and may give rise to a higher frequency of mutations causing albinism in certain groups in Africa.

Biological factors may also suggest heterozygous advantage. The tsetse fly has been found to settle more often on darker skin than lighter skin. Hence, in the Negroids, trypanosomiases ('sleeping sickness') is less likely to occur in the lighter pigmented heterozygotes than the darker homozygous normal individuals [Buxton, 1955].

1.2 Albinism and other disorders of hypopigmentation

Factors that affect the melanin biosynthetic pathway will have an effect on pigmentation. If an error occurs in this pathway by the lack of action of any of the components involved in it hypopigmentation will result. This can also occur if the process of transfer of the melanosomes from the melanocytes to the keratinocytes is interfered with [Robins, 1991]. Some of the possible hypopigmentation disorders that can arise are discussed below, with emphasis being placed on albinism. Genetic disorders of pigmentation can be classified into those of melanin production and disorders of melanocyte development.

1.2.1 Albinism: a disorder of melanin production

Albinism is found in all human populations, many animal populations, and even in the plant kingdom. It is a heritable group of disorders that occur due to malfunctions in the melanin biosynthetic pathway. These disorders are usually inherited as autosomal recessive traits and are

clinically characterised as a "congenital reduction or absence of melanin pigment in association with specific developmental changes in the optic nerve system resulting from the hypopigmentation" [Oetting *et al*, 1996]. Albinism can be divided into two main groups: ocular albinism (OA) and oculocutaneous albinism (OCA). The former is associated with the loss of pigmentation in the eyes alone, and the latter manifests as a lack of pigmentation in the skin, hair and eyes [Robins, 1991]. OCA is the commoner of the two. It is the most common autosomal recessive disorder in South African Negroid individuals [Manga *et al*, 1997].

1.2.1.1 The incidence and prevalence of albinism

The incidence of albinism differs greatly between populations. In Holland it has been reported to be 1/20 000 births, in Northern Ireland approximately 1/10 000, in Norway 1/9 600, and in Nigeria 1/5 000 [referenced in Kromberg and Jenkins, 1982]. The frequency of OCA, as a whole, is estimated to be 1/20 000 in most populations [King *et al*, 1995]. In South Africa, the prevalence among Bantu-speakers in Soweto was found to be 1/3 900 [Kromberg and Jenkins, 1982]. In the African American population the prevalence has been reported to be 1/10 000 and 1/18 000 in American Caucasians [King *et al*, 1995]. It appears that albinism is more common in populations of African origin than in European populations.

1.2.1.2 Clinical features of oculocutaneous albinism

The basic characteristics for clinical diagnosis of albinism are the absence or reduction of melanin pigment in the skin, hair and eyes. Other ocular features include hypoplasia of the fovea (maldevelopment), nystagmus (involuntary horizontal oscillations of the eyeball), photophobia (sensitivity to sunlight) and decreased visual acuity (in some cases partial blindness) [Witkop *et al*, 1983; Robins, 1991]. Misrouting of the optic fibres at the chiasm leads to the optical problems seen in individuals with albinism [Stevens *et al*, 1995]. The same eye defects are observed in the OA patients [Robins, 1991]. The relative reduction or lack of pigmentation differs between OCA individuals, and may lead to them being classified into different types.

1.2.1.3 Different types of albinism

Albinism can be divided into 6 main classes [King *et al*, 1995]. They are:

1. OCA1 or tyrosinase-related OCA, which can be subdivided into 4 groups:

a) OCA1A or tyrosinase-negative OCA
b) OCA1B or yellow OCA
c) OCA1MP or minimal pigment OCA
d) OCA1TS or temperature sensitive OCA

2. OCA2 or tyrosinase-positive OCA
3. Brown OCA (BOCA)
4. Rufous OCA (ROCA) or OCA3 [Manga *et al*, 1997]
5. Autosomal dominant OCA
6. Ocular albinism, which includes:
   a) OA1 (X-linked)
   b) Autosomal recessive OA (AROA)
   c) OA and deafness

The various OCA classes differ with regard to their main clinical features. These differences are shown in Table 1.2 and are used to classify the OCA subtypes.

1.2.1.4 The molecular basis of albinism

The albinism phenotype occurs because of changes that occur in the pigmentation system. These changes are caused by mutations that arise in the genes involved in pigmentation. The details of the genes and the proteins that result will be discussed in detail in section 1.3. However, in this subsection these genes will be mentioned with regard to albinism, to discuss their role and the findings that have arisen in connection with albinism.

OCA1A occurs because of mutations in the tyrosinase gene (*TYR*) that result in the lack of the enzyme in the melanocytes. Several different mutations have been found in various population groups that link OCA1 to *TYR* [Hearing and Jimenez, 1989; Giebel *et al*, 1991a; Spritz *et al*, 1991; King and Oetting, 1992; Tomita *et al*, 1992; Tripathi *et al*, 1993; King *et al*, 1995; Spritz and Hearing, 1995]. These mutations produce proteins that cannot function, and this results in a lack of melanin production. Hair-bulb tyrosinase activity is absent [King and Witkop, 1976; King *et al*, 1995]. This is measured using the *hair-bulb tyrosinase assay,* which was originally designed to differentiate between tyrosinase-positive and tyrosinase-negative OCA. The assay is

basically an enzymatic method of determining the activity of tyrosinase in hair-bulbs by exposing them to L-tyrosine at 37°C and observing the production of melanin in them [Vandorp *et al*, 1982]. In tyrosinase-negative OCA melanin is not produced, while in tyrosinase-positive OCA melanin is produced in the hair-bulbs.

OCA1B is also associated with *TYR* mutations that result in an enzyme with only some activity, which varies depending on the mutation and which have been called leaky mutations [Giebel *et al*, 1991b; King *et al*, 1995; Spritz and Hearing, 1995]. As the individual grows older, however, there is an accumulation of pigment, mainly red and yellow pigment [Giebel and Spritz, 1992; King *et al*, 1995]. Hair-bulb tyrosinase activity is low or absent in these individuals [King *et al*, 1995]. This type of albinism was first discovered in the Amish community [Giebel *et al*, 1991b].

OCA1MP is associated with the absence of skin, hair and eye pigmentation at birth and the formation of small amounts of iris pigment in the first decade of life. Mutations in the tyrosinase gene have been found to produce this phenotype. Hair-bulb tyrosinase activity is absent in these individuals [King *et al*, 1986, 1995]

OCA1TS is associated with *TYR* mutations that produce a protein which is temperature sensitive and loses activity above 35°C [King *et al*, 1995; Oetting *et al*, 1996]. Hence, melanin synthesis occurs in the cooler areas of the body, such as the arms and legs. The hair-bulb melanocyte and melanosome architecture is normal [King *et al*, 1995]. This phenotype has been compared to the Siamese cat and Himalayan mouse [Giebel *et al*, 1991c; Oetting *et al*, 1996].

OCA2 individuals have reduced pigmentation of the skin, hair and eyes [Stevens *et al*, 1995]. The hair is yellow at birth in African Americans and Africans with OCA2, and remains yellow throughout life or may turn darker [King *et al*, 1995]. Hair-bulb tyrosinase activity is normal or nearly normal and the tyrosinase protein is normal [King and Witkop, 1976; King *et al*, 1995]. This is the most common type in southern Africa [Lund *et al*, 1997] and has been found in Tanzania, Zimbabwe, South Africa and other parts of sub-Saharan Africa [Oetting *et al*, 1996; Lund *et al*, 1997]. OCA2 is very common in African-Americans with a frequency of 1 in 10 000 [King *et al*, 1995; Durham-Pierre *et al*, 1996]. It has a frequency of approximately 1 in 36 000 in

the American Caucasoids [King *et al*, 1995; Oetting *et al*, 1996]. The prevalence of OCA2 has been reported as 1/3 900 in Bantu-speaking Negroid individuals in South Africa [Kromberg *et al*, 1982]. The *P* gene has been found to be associated with this type of albinism [Ramsay *et al*, 1992; Rinchik *et al*, 1993]. A 2.7kb deletion removes the whole of exon 7 of this gene [King *et al*, 1995; Durham-Pierre *et al*, 1996; Oetting *et al*, 1996; Lund *et al*, 1997], which ultimately results in a frameshift and truncation of the predicted polypeptide [Durham-Pierre *et al*, 1996] Point mutations in the *P* gene have also been associated with OCA2 [Lee *et al*, 1995].

BOCA and ROCA individuals are difficult to diagnose because of their similar pigmented appearances. Since many of these individuals have the ocular features of albinism they have been classified as OCA types [King *et al*, 1995]. In BOCA Africans and African-Americans the hair and skin colour is brown, which may grow darker with time [King *et al*, 1995; King *et al*, 1995]. In Caucasoid individuals the hair is golden blond and the skin is pale. The hair-bulb tyrosinase activity is normal. The melanosomes have very little melanin, suggesting that they arrest during stages I and II of melanosome development [King *et al*, 1985]. There is some confusion with regard to the gene that causes BOCA. According to Boissy *et al* (1996) a mutation of the *TYRP1* gene, the 368delA mutation in exon 6, was thought to be responsible for the BOCA phenotype as it was found in the homozygous state in one so called "BOCA infant". *TYRP1* may be responsible for the production of the DHICA oxidase enzyme in the melanin biosynthetic pathway. A study by Manga *et al* (1997) has shown, however, that this mutation causes the ROCA phenotype in African individuals. It is possible that Boissy's group evaluated the infant too early (at one day) and described him as having BOCA. The phenotype may, however, change with age and he may actually have ROCA [Oetting *et al*, 1996]. Alternatively, the genetic background (mixed ancestry) in the individual may modify the phenotype. ROCA individuals have red-bronze skin colour, ginger red hair, and blue or brown irides. They have been reported in African individuals, as well as in individuals in Papua New Guinea [Walsh, 1971]. Another mutation found to be associated with ROCA is the S166X nonsense mutation in exon 3 of *TYRP1* gene [Manga *et al*, 1997]. The two mutations accounted for 95% of the ROCA mutations in the study by Manga *et al* (1997). The prevalence of BOCA in southern Africa is unknown, but that of ROCA is approximately 1/8 500 [Kromberg *et al*, 1990, cited in Manga *et al*, 1997]. The prevalence of BOCA in Nigeria has been reported as being 1/10 000 [King *et al*, 1995]. One

family studied by Manga *et al* (1997) had siblings with a clinically unclassified form of albinism, who were found to be compound heterozygous for the 368delA and S166X mutations in the *TYRP1* gene. These two mutations generate truncated forms of the protein. The individuals were also heterozygous for the 2.7kb deletion in the *P* gene, suggesting that the phenotype may be due to some interaction between the loci.

Autosomal dominant OCA is the only type that has a dominant pattern of inheritance. In general these individuals have decreased pigmentation [King *et al*, 1995]. It is uncertain what the molecular characterisation of this OCA type is.

Ocular albinism results in normally pigmented hair and skin, and light eyes. OA1 is thought to be associated with mutations in the *OA1* gene on the X chromosome [Bassi *et al*, 1994]. It is found close to the pseudoautosomal region of the X chromosome at position Xp22.3 [King *et al*, 1995; Newton *et al*, 1996; Oetting *et al*, 1996]. The function of the protein this gene encodes is not known [King *et al*, 1995; Newton *et al*, 1996]. Other genes on the X chromosome may also be involved in this disorder, as not all patients with OA1 have mutations in this gene [King *et al*, 1995]. AROA has a similar phenotype to OA1. Although the main gene responsible for this phenotype has not been identified, a study by Fukai *et al* (1995a) identified *TYR* mutations in two AROA individuals. The *P* gene has also been associated with a form of OA. The third OA type is rare and has clinical features similar to OA1 in addition to deafness and has been seen in an Afrikaner family [King *et al*, 1995]. It has not been investigated as extensively as the other OAs and not much is known about its molecular defect.

1.2.1.5 Albinism in South Africa

In South African Blacks the most frequent types of albinism are OCA2, BOCA and ROCA [Kromberg *et al*, 1987]; with OCA2 being the commonest of the three [Kromberg, 1987]. Albinism has a prevalence of 1/3 900 in South African Bantu-speakers, but the prevalence varies between the different South African ethnic groups [Kromberg and Jenkins, 1982]. The highest rates were approximately 1 in 2 000 in the Sotho-Tswana and Swazi and lower in the Zulu and Xhosa groups (1 in 4 500) [Kromberg, 1987].

Table 1.2: The characteristics of the main types of OCA [Bolognia and Pawelek, 1988; King *et al*, 1995]

Characteristic	OCA1A	OCA1B	OCA1MP	OCA1TS	OCA2	ROCA	BOCA
Skin colour	Pink-white	Pink-white	Pink-white	White with pigmented arms and legs	Pink-white to cream	Reddish brown	Light brown
Hair colour	White through life	White at birth, turns yellow with age	White at birth, yellow tint with time	White at birth, changes after puberty. Arms and legs pigment with time	Pigmented at birth, darkens with age	Reddish brown to mahogany	Light brown
Iris colour	Blue to grey	Blue to grey	Blue to grey	Blue to grey	Blue to hazel to tan	Hazel to reddish brown	Blue grey to light brown
Skin tanning	Absent	Possible	Absent	Unknown	Absent	Possible	Possible
Nystagmus	Present	Present	Present	Present	Present	Present or absent	Present
Photophobia	Present	Present	Present	Present	Present	Mild to none	Mild to None
Hair-bulb melanosomes	Stage I and II	Stage I, II and III	Stage I, II and III	Scalp and axilla stage I and II; arms and legs stages I, II and III	Stage II and III	Stage III and IV	Stage II and III, and a few stage IV
Hair-bulb tyrosinase activity	Absent	Absent	Absent	Absent	Normal	?	Normal
Gene involved	*TYR*	*TYR*	*TYR*	*TYR*	*P*	Possibly *TYRP1*	*TYRP1*

1.2.2 Other hypopigmentation disorders

A number of other hypopigmentary disorders have albinism or reduced pigmentation in addition to their characteristic clinical features, where the primary defect is not limited to the melanocyte. These include *Prader-Willi* and *Angelman* syndromes, which have been associated with OCA2; and *Chadiak-Higashi* and *Hermansky-Pudlak* syndromes, which have been described as types of albinism [King *et al*, 1995; Orlow, 1995]. These and other disorders of hypopigmentation will be discussed briefly, with emphasis on their molecular defects.

1.2.2.1 Prader-Willi syndrome (PWS)

PWS is a developmental syndrome that has been found to be associated with OCA2 or hypopigmentation. This disorder occurs at a frequency of 1/10 000 to 1/25 000 [Butler, 1989, cited in Glenn, 1997]. It is characterised by neonatal hypotonia and failure to thrive, hyperphagia leading to obesity, hypogonadism, small hands and feet, and mental retardation with a characteristic behavioural disorder [reviewed in King *et al*, 1995; Gunay-Aygun *et al*, 1997]. Although almost half the affected individuals have hypopigmentation of the skin, hair and eyes; most do not have typical ocular features of albinism. Nystagmus and strabismus are common but transient and the fovea are not entirely normal although foveal hypoplasia is not present. Misrouting of the optic tracts has been observed in some PWS individuals with hypopigmentation. Hair-bulb tyrosinase activity is also low [reviewed in King *et al*, 1995].

PWS is caused by the lack of expression of imprinted gene(s) on the paternal chromosome 15q11-13. This mode of inheritance can be said to occur because the male and female chromosomal regions, in the germline, are "marked" differently [Gunay-Aygun *et al*, 1997]. Approximately 70% of PWS individuals have a deletion on the long arm of the paternal chromosome 15 and most of the others have uniparental disomy of the maternal chromosome 15 [reviewed in King *et al*, 1995]. In both cases information is absent from the paternal chromosome 15. Some PWS patients (about 5%) that do not have either of the above mutations may have abnormal maternal DNA methylation imprints in the PWS region [Gunay-Aygun *et al*, 1997]. A study by Lee *et al* (1994) presented a patient with OCA2 and PWS who had a deletion in the paternal chromosome 15 and a mutant *P* allele on the maternally inherited chromosome 15. The hypopigmentation phenotype of PWS patients may be due to the deletion of the *P* gene

[Spritz *et al*, 1997]. Rinchik *et al* (1993) suggested the possible hemizygosity of the *P* gene to be the cause of hypopigmentation, with the remaining allele being functionally deficient. However, Spritz and colleagues (1997) obtained results that showed multiple *P* gene haplotypes in hypopigmented PWS individuals, perhaps arguing against the above hypothesis.

1.2.2.2 Angelman syndrome (AS)

AS is also a developmental disorder with "developmental delay and severe mental retardation, microcephaly, neonatal hypotonia, ataxic movements and inappropriate laughter" [reviewed in King *et al*, 1995]. It is also associated with hypopigmentation of the skin, hair and eyes. There is low hair-bulb tyrosinase activity and incomplete melanisation of the melanosome [King *et al*, 1993]. These individuals may have a history of nystagmus or strabismus and reduced retinal pigmentation [reviewed in King *et al*, 1995]. AS has an incidence of 1 in 20 000 births [Chan *et al*, 1993].

As with PWS, genetic imprinting characterises the inheritance of this disease. AS arises from a lack of maternal contribution of the 15q11-13 chromosomal region; which may have occurred as a result of deletion, in most cases, or uniparental disomy, in some [Chan *et al*, 1993]. Others may have paternal DNA methylation on maternally inherited chromosome 15q11-13 [Gunay-Aygun *et al*, 1997]. Hypopigmentation is often found in deletion cases [Butler, 1989, cited in Spritz *et al*, 1997].

1.2.2.3 Hermansky-Pudlak syndrome (HPS)

HPS is a multi-system disorder. The patients with HPS present with OCA similar to that described above for OCA1 and OCA2 individuals [Oetting *et al*, 1996]. The patients have light blond to reddish-brown hair and light skin [Orlow, 1995]. Iris colour varies from blue to brown and all the common ocular features of albinism are present. Hair-bulb tyrosinase activity is present but low. The OCA arises as a result of an unknown primary defect affecting the melanocytes, platelets and possibly other tissues [King *et al*, 1995]. In addition they have a mild diathesis and a ceroid disease affecting the lungs (pulmonary) and gut (intestinal) [Fukai *et al*, 1995b; King *et al*, 1995]. Ceroid lipofuscin, a yellow-brown pigment, accumulates in macrophages in the lung, resulting in the lung problems, and in the gut [Orlow, 1995]. A

bleeding disorder also occurs due to defective platelet aggregation. The average survival age is 30-50 years. HPS has a high frequency in Puerto Rico [King *et al*, 1995; Orlow, 1995], where it occurs in 1 in 2 700 individuals [Oetting *et al*, 1996]. It is also frequent in an isolated village in the Swiss Alps [Fukai *et al*, 1995b]. However, it has a low frequency in most other populations [Fukai *et al*, 1995b; Oetting *et al*, 1996].

The melanocytes and melanosome ultrastructures are normal in individuals with HPS [King *et al*, 1995]. HPS platelets are deficient in the protein granulophysin. This disorder is inherited as an autosomal recessive trait. The *HPS* gene was mapped to chromosome 10q24-25. Mutation analysis confirmed the founder effect in the Puerto Rican population and the population of the Swiss Alps village [Fukai *et al*, 1995b; Oetting *et al*, 1996]. There is some evidence, however, for locus heterogeneity [Hazelwood *et al*, 1997; Oh *et al*, 1998], as studies on patients from Puerto Rico and other geographical locations have shown that not all HPS individuals had mutations in the *HPS* gene.

1.2.2.4 Chediak-Higashi syndrome (CHS)

CHS has been described as a disease with recurrent pyogenic sinopulmonary and cutaneous infections [Bolognia and Paweleck, 1988]. CHS is another type of incomplete albinism. Skin, hair and eye pigmentation is reduced [King *et al*, 1995]. Hair colour is light brown to blond, and has a metallic silver-grey sheen. The skin is creamy-white to slate grey. Nystagmus and photophobia may be present or absent. Iris pigment is reduced and the misrouting of the optic fibres have been described as being similar to OCA1 and OCA2 individuals. Patients have an increased susceptibility to bacterial infections and giant peroxidase-positive lysosomal granules in peripheral blood granulocytes [King *et al*, 1995]. The granulocytes cannot discharge their contents into phagocytic vacuoles normally. The final outcome for CHS patients is death at an early age [Bolognia and Paweleck, 1988].

Giant melanosomes form in melanocytes, which cannot be transferred into the keratinocytes. Hence, there is abnormal melanosome distribution and hypopigmentation. The hair-shaft has large and irregular pigment granules, this feature is used for prenatal diagnosis [King *et al*,

1995]. CHS has an autosomal recessive pattern of inheritance. The gene for CHS in humans, *CHS1*, has been mapped to chromosome region 1q42.1-q42.2 [Barrat *et al*, 1996].

1.2.2.5 Piebaldism

Piebaldism has been referred to as partial albinism [Bolognia and Paweleck, 1988], and has also been described as a disorder of melanocyte development [Spritz and Hearing, 1995]. Patches of white skin are found on the forehead, chest, abdomen, and the mid-portion of the upper and lower extremities. A white forelock (*poliosis*) may also be present. The depigmented areas have a general absence of melanocytes or at least a considerable decrease in their numbers [Orlow, 1995]. This disorder has an autosomal dominant pattern of inheritance. The gene involved is the *KIT* gene, which encodes a growth factor receptor that resides on the surface of the melanocytes. Its function is to inform the melanocyte about growth and migration. The mutations found in the gene disrupt the function of this receptor. Thus, the distribution of the pigment cells is affected during embryo development [Orlow, 1995].

1.2.2.6 Waardenburg's syndrome (WS)

Waardenburg's syndrome is another hypopigmentary disorder of melanocyte development. Three types of WS are known, types I to III, of which type I is the classic type [Spritz and Hearing, 1995]. These patients have a depigmented patch, most often V-shaped, on the central forehead in association with poliosis [Orlow, 1995]. Heterochromic irides (different coloured irises) may be present and depigmented patches on the trunk and extremities are rare. Sensorineural deafness is often present. Other features are dystopia canthorum, confluence of the medial eyebrows, a broadened nasal root and thin upper lip [King *et al*, 1995; Orlow, 1995]. Type II lacks dystopia canthorum and type III is associated with limb abnormalities in addition to the type I phenotype [Spritz and Hearing, 1995].

The melanocytes of the iris, skin and hair are involved. The disorder has an autosomal dominant pattern of inheritance [Orlow, 1995]. WS I and II have been associated with mutations in the *PAX3* gene on the distal long arm of chromosome 2 [Baldwin *et al*, 1995; Spritz and Hearing, 1995]. This gene encodes a transcription factor necessary for the differentiation of neural crest cells. The neural crest gives rise to melanocytes in most areas of the body and to other

centrofacial structures [Orlow, 1995], which explains why mutations in this gene give rise to many of the clinical features. Type II has been thought to be associated with the *MITF* gene, the human homologue of the mouse *microphthalmia* gene, located on the short arm of chromosome 3 [Spritz and Hearing, 1995].

1.2.2.7 Vitiligo

Vitiligo is another disorder of melanocyte function [Bolognia and Pawelek, 1988]. Two types of vitiligo exist- type A, the most common, and type B. The depigmented areas in vitiligo are: periorifical (around the mouth, nose, eyes, nipples, umbilicus, and anus), flexor wrists, extensor surface of the extremities (such as the elbows and knees), intertriginous zones (groin and axilla), and oral mucosa. The depigmentation is a result of a loss of melanocytes. Type A has the typical features as described above, and type B has depigmentation that is unilateral and dermatomal-like in distribution. Several autoimmune disorders have also been associated with this disorder, including hyperthyroidism and IDDM (insulin dependent diabetes mellitus) [Bolognia and Pawelek, 1988]. The melanocytes in the depigmented areas are destroyed because of autoantibodies and other forces of destruction. There is evidence that the melanocyte destruction is reversible early in the disease [Norris *et al*, 1994]. Repigmentation has been found to occur when pigmented islands develop, which expand and coalesce. These islands have been shown to occur from amelanotic melanocytes in the outer root sheath of hair follicles, which migrate into the epidermis and function as normal epidermal melanocytes [reviewed in Morelli *et al*, 1993]. By 1988 the mode of inheritance was thought to be either polygenic or autosomal dominant [Bolognia and Pawelek, 1988]. To date there seems to be no clear linkage to any particular genetic locus.

1.3 Genes in pigmentation

More than 60 loci have been found to regulate pigmentation in mice. The α-MSH and agouti proteins act outside the cell, the MC1R acts on the membrane of the cell, and the enzymes tyrosinase, DOPAchrome tautomerase, and DHICA oxidase work within the cell in the melanin biosynthetic pathway. The genes that encode each of these have been studied extensively in the

mouse, setting the scene for much of the research on human genes. These proteins and the genes that encode them will be discussed in detail.

1.3.1 Genes controlling melanin synthesis

The two main biological components that stimulate melanogenesis are α-MSH and MC1R. They will be discussed together with the agouti protein, which may also be a part of the stimulatory mechanism in humans.

1.3.1.1 The α-melanocyte stimulating hormone locus

The *α-MSH* gene, or the *POMC* gene, is on the short arm of chromosome 2 in human. The mouse *POMC* gene is located on chromosome 12 [Uhler *et al*, 1983]. The gene encodes a large precursor peptide called POMC (29 500 Mr [Eberle, 1980]), which is produced in the pituitary gland and is processed into nine different peptides (Fig.1.12) that have various functions [Chang *et al*, 1980; Mountjoy *et al*, 1992]. The peptide is first processed to *adrenocorticotropin* (ACTH) and *β-lipotropin* (β-LPH) [Cochet *et al*, 1982; Mountjoy *et al*, 1992]. ACTH is then further enzymatically cleaved, in the intermediate lobe of the pituitary, to α-MSH and *corticotrophin-like intermediate lobe peptide* (CLIP) [Chang *et al*, 1980; Eberle, 1980]. α-MSH is released from the pituitary into the blood [Valverde *et al*, 1996]. In addition to its functions in pigmentation, other functions of α-MSH include temperature control in the septal region of the brain and stimulation of prolactin release from the pituitary [Mountjoy *et al*, 1992]. The *POMC* gene is composed of three exons, the third exon of which is responsible for the production of ACTH, and thus α-MSH. An elongation or shortening of this 13 amino acid hormone leads to a decrease in its activity [Eberle, 1980; Grantz *et al*, 1994; Wilson *et al*, 1995].

α-MSH induces cell proliferation and increases DNA synthesis [Eberle, 1980]. It stimulates proliferation by increasing the entry of the melanocytes into S-phase of the cell cycle [Abdel-Malek *et al*, 1994]. UV radiation increases MSH and/or POMC levels in human plasma cells and keratinocytes [Chakraborty and Pawelek, 1993], which results in increased eumelanogenesis and pigmentation or tanning. This has been supported by many studies, one of which has shown that

if α-MSH is injected into the skin of dominant agouti mutant yellow mice (A^y) the immediate melanocytes respond by producing eumelanin [Jackson, 1993a].

Mutations in the *POMC* gene involving *α-MSH* have been identified in a study by Krude *et al* (1998), which reported findings on two individuals that had the following symptoms: obesity, ACTH deficiency and alteration of pigmentation (red hair). One patient was homozygous for a C3804A mutation in exon 2 and the second patient had two mutations in exon 3, G7013T and C7133Δ, which interfere with ACTH and α-MSH synthesis. Therefore, there seems to be a link between the *α-MSH* gene and red hair, a normal variation in hair colour. This is a point of interest for studies on individuals with red hair.

1.3.1.2 The melanocyte stimulating hormone receptor locus (*MC1R*)

The *MC1R* gene has been cloned [Mountjoy *et al*, 1992] and mapped to human chromosome 16 at locus 16q24.3 [Grantz *et al*, 1994; Magenis *et al*, 1994]. The mouse homologue *e* (the extension locus) is found on mouse chromosome 8 (Table 1.3) [Magenis *et al*, 1994]. The human *MCIR* gene is 951 bases long and comprises only one exon, which encodes a 317 amino acid protein [Rana *et al*, 1999].

Animal studies carried out by many groups have linked mutations at the *e* locus to variation in animal coat colour. In mice, five different alleles have been described that give rise to various coat colours: the wild-type (E^+), recessive yellow (*e*), tobacco darkening (E^{tob}), sombre (E^{so}), and sombre-3J ($E^{so\text{-}3J}$). All the alleles listed are dominant to *e*. The *e/e* mouse is almost entirely yellow because of the absence of eumelanin in the hair follicles. $E^{so\text{-}3J}$ is phenotypically similar to the E^{so} mice [Robbins *et al*, 1993]. E^{tob} produces a darkening of the dorsum with agouti flanks. E^{so} homozygotes are mostly black with a few yellow hairs and darkened skin, resembling extreme nonagouti mice. Therefore, all the dominant gain-of-function alleles result in black hairs while the recessive result in yellow [Jackson, 1993a]. The dominant E^{tob} and E^{so} alleles have an epistatic effect on the *agouti* locus, as they prevent the production of red pigmentation by the agouti protein. A recent study by Vage *et al* (1997), however, revealed a dominant allele at the *e* locus in foxes, which resulted in red coat colour, therefore, an apparent non-epistatic interaction between the agouti locus and the dominant extension allele.

The coat colour of the yellow Labrador retriever and the red fox occur because of recessive alleles at the extension locus [Koer and Pawelek, 1982]. Dominant extension alleles are also thought to be responsible for the black masking in the Great Dane and German shepherd. Similar relationships have been indicated in the rabbit and guinea pig, where dominant and recessive alleles have been observed [Koer and Pawelek, 1982]. A recent study by Rana *et al* (1999) on primates; such as the pigmy and common chimpanzee, gorilla, orangutan, and baboon; revealed only a few differences in the amino acid sequence to the human sequence. " A parsimony tree indicated that the ancestral MC1R protein sequence can be inferred to be identical to that of the human consensus sequence" [Rana *et al*, 1999]. The *MC1R* of the gorilla, chimpanzee and human appear to have evolved faster than that of the baboon and orangutan.

Many studies have found *MC1R* gene variation to be linked to the red hair and pale skin phenotypes in humans [Valverde *et al*, 1995; Box *et al*, 1997; Smith *et al*, 1998; Rana *et al*, 1999]. This gene has also been studied in other normally pigmented individuals in different geographical areas of the world [Box *et al*, 1997; Koppula *et al*, 1997; Rana *et al*, 1999]. The results of these studies will be discussed in detail in section 1.4.1.

Since, variation in the *MC1R* gene is associated with the red hair and sun-sensitive skin it was considered a candidate gene for melanoma susceptibility. A study by Valverde *et al* (1996) has shown that *MC1R* variants are more common in melanoma cases than controls. The variant D84E was only present in melanoma cases. Thus, this gene in humans seems to have a "casual" link to melanoma occurrence [Valverde *et al*, 1996]. Cultured murine melanoma cells were found to express more MSH binding sites in response to UV light, which implies that MC1R transduces the signal for the melanogenic effects of UV light [Jackson *et al*, 1994a]. UV radiation was also found to increase the activity of these receptors in mouse melanoma cells [Chakraborty and Pawelek, 1993].

1.3.1.3 The agouti locus

The *agouti* locus (*a*) in mice encodes the agouti signalling protein (ASIP or ASP), whose action in the mouse is thought to be as an antagonist to α-MSH action. ASP is produced by dermal papillae cells and modulates the production of pigment by follicular melanocytes [Sakai *et al*,

1997]. The mouse *a* locus (Table 1.3), on chromosome 2, encodes a 131 amino acid protein [Bultman *et al*, 1992]. The human homologue, called the *ASP* gene, was found on chromosome 20 [Wilson *et al*, 1995] and encodes a protein that is 132 amino acids in length. It has 80% identity to the mouse homologue [Suzuki *et al*, 1997].

In mice the *a* allele results in the production of eumelanin in the first four days of the hair growth cycle, followed by the production of ASP during days 4 to 6 causing phaeomelanin to be produced. After 6 days the agouti gene is turned off again and eumelanin is produced again. This mechanism results in the stripe of yellow on the black hair of mice described earlier [Sakai *et al*, 1997]. Mutations in the *a* locus may lead to the over-expression and hyperfunction of ASP or the under-expression and non-function of it. At least 18 alleles have been identified in mice, the combinations of which give rise to subtle differences in coat colour [Bultman *et al*, 1992]. The dominant lethal yellow mutation (A^y) results in completely yellow hairs when in heterozygous form [Bultman *et al*, 1992; Sakai *et al*, 1997], while the homozygous form is lethal [King *et al*, 1995]. This allele and the viable yellow allele (A^{vy}) have been linked to obesity and diabetes mellitus in mice [Wilson *et al*, 1995]. The recessive non-agouti allele (a) gives rise to hair that is mostly black and the extreme non-agouti allele (a^e) produces completely black hairs [Bultman *et al*, 1992; Sakai *et al*, 1997]. The black-tan allele (a^t) gives rise to an all black dorsum and an all yellow ventrum [Bultman *et al*, 1992].

In humans it is unclear how the agouti homologue functions and whether it plays a role in phaeomelanin production. As mentioned earlier, Suzuki *et al*'s (1997) study has shown that the agouti signalling protein blocked the binding of α-MSH to MC1R and, therefore, inhibited the effects of α-MSH in human melanocytes. ASP was found to inhibit basal levels of tyrosinase activity. The agouti protein has also been shown to markedly inhibit the production of eumelanosomes in melan-a murine melanocytes [Sakai *et al*, 1997]. Thus, the murine and human ASP seems to have similar roles.

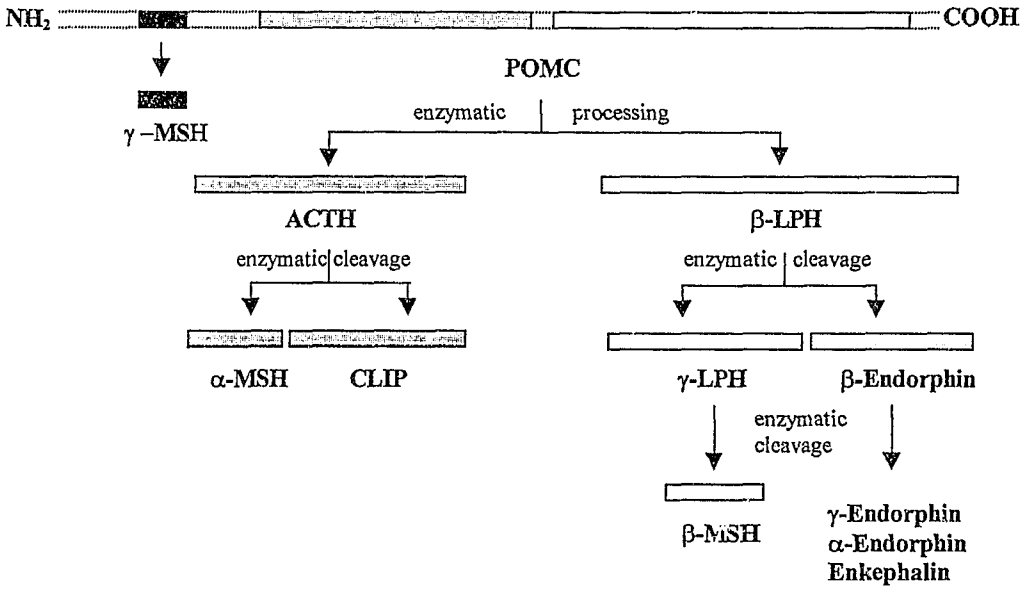


Fig.1.12: The products of the *POMC* locus. The cleaved products of the POMC protein. [Figure modified from Eberle, 1980; Bolognia Pawelek, 1988].

Table 1.3: Genes involved in melanogenesis

Protein	Gene	Locus	Gene product	Mouse locus
Tyrosinase	*TYR*	11q14-21	Tyrosinase	albino (*c*)
Tyrosinase related protein 1	*TYRP1*	9p23	Believed to be DHICAoxidase	brown (*b*)
Tyrosinase related protein 2	*TYRP2*	13q32	DOPAchrome Tautomerase	slaty (*Slt*)
α-Melanocyte stimulating hormone	*POMC/ACTH/ α-MSH*	2p23	α-Melanocyte stimulating hormone	*POMC*
Melanocortin 1 receptor	*MSH-R/MC1R*	16q24.3	Melanocyte stimulating hormone receptor	extension (*e*)
Agouti	*ASP*	20q11.2	Agouti signalling protein	agouti (*a*)
P	*P*	15q11.2-12	Uncertain	pink-eyed dilution (*p*)

1.3.2 Genes involved in the melanin biosynthetic pathway

The melanin biosynthetic pathway involves many components, including enzymes and other proteins essential for pigment production, without which the process would not take place or would occur at a much slower rate. The genes for many of these components have been identified and cloned. One of the essential enzymes is the tyrosinase enzyme, without which melanin biosynthesis cannot occur. The tyrosinase (*TYR*) gene and the genes for DHICA oxidase (*TYRP1*) and DOPAchrome tautomerase (*TYRP2*) will be discussed, as well as the *P* gene, whose function is still uncertain.

1.3.2.1 The tyrosinase related protein (TYRP) family

The tyrosinase related family of proteins is made up of tyrosinase, DOPAchrome tautomerase, and DHICA oxidase encoded by the *TYR*, *TYRP1* and *TYRP2* genes respectively in humans. These loci are homologous to the albino (*c*), brown (*b*), and slaty (*slt*) loci in the mouse (Table 1.3) and encode proteins with similar structures and features, but with distinct catalytic capacities [Winder *et al*, 1994]. The proteins encoded by these genes have the following common features: an N-terminal growth factor-like repeat domain, two copper binding sites, a cystein rich region, and a C-terminal transmembrane domain (Fig.1.13) [Jackson *et al*, 1991; Jackson *et al*, 1992; Morrison *et al*, 1995; Sturm *et al*, 1995]. They also have a common C-terminal membrane spanning exon [Sturm *et al*, 1995]. The transmembrane domain is the one by which the protein is anchored to the inner surface of the melanosomal membrane [Jackson *et al*, 1991].

The three genes are thought to have evolved from one gene. The hypothesis is that *TYR* gave rise to *TYRP1* by duplication, which in turn duplicated to give rise to the *TYRP2* gene [Sturm *et al*, 1995]. The strong similarity between vertebrate *TYRP1* and *TYRP2* sequences, as opposed to either one with the *TYR* sequence suggests that the tyrosinase-related proteins diverged from tyrosinase before duplicating [Jackson *et al*, 1994b]. The sequence similarity between *TYR* and *TYRP1* indicated an evolutionary ancestor [Jackson *et al*, 1991]. At the amino acid level, *TYR* and *TYRP1* have 43% homology [Orlow *et al*, 1993] and 55.3% nucleotide sequence homology [Murty *et al*, 1992]. The exonic structures in the mouse and human *TYR* and *TYRP1* genes are very different.

The position of some introns are conserved, for example, human *TYR* introns II and IV have homologous positions to those of *TYRP1* introns V and VII, if considering the first non-coding exon of *TYRP1* (introns IV and VI if not). The *TYRP1* introns III and VII (or II and VII, for the reasons given above) have homologous positions to the *TYRP2* introns III and VII [Jackson *et al*, 1991; Sturm *et al*, 1995] (Fig.1.13). *TYR* and *TYRP1* lack promoter similarity [Jackson *et al*, 1991].

1.3.2.1.1 The tyrosinase locus

The *TYR* gene is located at 11q14-21 and is the homologue of the mouse *albino* (*c*) locus on mouse chromosome 7 [Barton *et al*, 1988; King *et al*, 1995]. The human *TYR* gene contains five exons and four introns [Jackson *et al*, 1991; King *et al*, 1995]. It encodes a 529 amino acid polypeptide [Sturm *et al*, 1995]. The nascent protein product is approximately 60kDa, which is glycosylated to produce a 75kDa product [Hearing and Tsukamoto, 1991; King *et al*, 1995]. The tyrosinase related gene (*TYRL*) is a site of tyrosinase cDNA hybridisation at position 11p11.2-cen, which consists of just exons 4 and 5 of this gene with the entire fourth and part of the third intron [Giebel *et al*, 1991d; King *et al*, 1995]. There is 98% homology between exons 4 and 5 of the *TYR* gene and *TYRL*. Since no *TYRL* transcription product is detectable in human melanocytes, it has been called a pseudogene [King *et al*, 1995].

Some mutations in the mouse and human tyrosinase loci result in the complete absence of pigmentation and OCA1 [Barsh, 1996]. Over 60 different mutations have been identified in the *TYR* gene in association with OCA1, which occurs in approximately 1 in 40 000 individuals in most populations [Oetting *et al*, 1996]. OCA1 includes tyrosinase-negative (OCA1A), yellow (OCA1B), minimal pigment (OCA1MP) and temperature sensitive (OCA1TS) OCA.

In mice, two *c* locus mutations result in the classic albino phenotype. Other alleles known include the *Himalayan* (c^h), which alters the glycosylation of tyrosinase resulting in a temperature sensitive form of the protein; and the *chinchilla* (c^{ch}), which increase the sensitivity of tyrosinase to proteolytic inactivation causing decrease in enzyme function [Hearing and Tsukamoto, 1991].

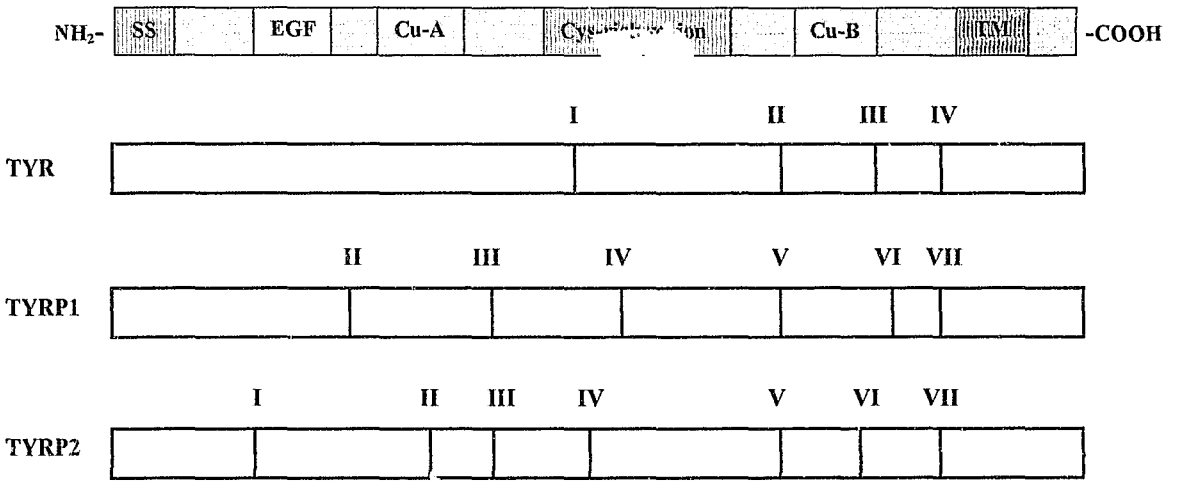


Fig.1.13: The TYRP family of proteins. This figure shows the relative positions of the introns (roman numerals) of each gene (*TYR*, *TYRP1* and *TYRP2*) of the TYRP family relative to the protein structure (top most). The positions of the N-terminal secretery signal peptide (SS), the epidermal growth factor-like region (EGF), two copper binding sites (Cu-A and Cu-B), the cystein rich region and the C-terminal transmembrane domain (TM) are shown. [Figure from Sturm *et al*, 1995].

Tyrosinase is a copper containing enzyme [Lerner *et al*, 1949] that plays a vital role in the melanin biosynthetic pathway. It takes part in three reactions, which have already been discussed. The tyrosinase hydroxylase activity of tyrosinase is thermostable, but its activities as DOPA oxidase and DHI oxidase are temperature sensitive [Tripathi *et al*, 1992a]. The copper in the tyrosinase molecule is essential for its activity, as studies have shown that copper-free tyrosinase has little activity [Lerner *et al*, 1950]. The copper residue has also been said to bind to tyrosine at the beginning of melanogenesis [Oetting and King, 1994].

1.3.2.1.2 The tyrosinase-related protein 1 locus

The human *TYRP1* gene has 93% homology, at the amino acid level, to the mouse *brown* (*b*) locus on mouse chromosome 4 [Cohen *et al*, 1990, cited in King *et al*, 1995; reviewed in King *et*

al, 1995]. The human *TYRP1* was assigned to chromosome 9p23 by *in situ* hybridisation [Murty *et al*, 1992]. This gene has 8 exons, the "first exon" being non-coding [Sturm *et al*, 1995] (Fig.1.13). It encodes a 527 amino acid polypeptide that is 60kDa in molecular weight. When mature the protein is a 75kDa trans-membrane melanomosomal glycoprotein, which is the human melanoma autoantigenic glycoprotein 75 (gp75), not present in non-pigmented melanoma cells [Hearing and Tsukamoto, 1991].

In humans, the S166X and 368delA mutations in exons 3 and 6 respectively result in the ROCA phenotype [Manga *et al*, 1997]. Mutations in the mouse *b* gene produce a brown coat colour rather than the wild-type black. The *cordovan-Harwell* allele (b^c) results in intermediate levels of pigment and is associated with very low levels of *Tyrp1* mRNA. The *white-base brown* (B^w) and the *light* allele (B^{lt}) are dominant. B^w results in the transcription of genes, which encode proteins that are toxic to melanocytes, resulting in melanocyte death. B^{lt} is a neomorph, which is a mutation that confers a new function to the protein, which apparently disrupts melanosome structure [King *et al*, 1995].

The function of the TYRP1 protein in humans and mice has been an issue of debate. Tyrp1 was found to act after the action of DOPAchrome tautomerase and could oxidise DHICA in the melanin biosynthetic pathway in mouse cells [Kobayashi *et al*, 1994a], although this may not be the case in humans [Boissy *et al*, 1998]. Kobayashi *et al* (1994b) found that in mouse melan-a (black), melan-b (brown), melan-c (albino) melanocytes and mouse melanoma cells *b* encoded a DHICA oxidase, and this function was confirmed by Jimenez-Cervantes *et al* (1994).

The protein was associated with eumelanin synthesis in mouse melanocytes [Bennette, 1991, cited in Del Marmol *et al*, 1993]. Many other functions have been ascribed to mouse Tyrp1 including catalase activity [Halaban and Moellmann, 1990], DOPAchrome tautomerase activity [Winder *et al*, 1993, cited in Box *et al*, 1998], as a DHI conversion factor [Pawelek, 1991, cited in Del Marmol *et al*, 1993], and another tyrosinase [Jimenez *et al*, 1991].

The TYRP1 protein in human melanocytes was found to be associated exclusively with the eumelanogenic pathway because *TYRP1* mRNA has been found only in eumelanin containing cells [Del Marmol *et al*, 1993]. In contrast to the mouse studies, melanocytes from a patient with OCA3, who lacks *TYRP1* gene expression, had the same DHICA oxidase activity as the

controls. They also found that expression of *TYRP1* in fibroblasts transfected with the sense *TYRP1* cDNA had the same DHICA oxidase activity as the controls [Boissy *et al*, 1998]. Since the mouse and human studies were done using different techniques they are not directly comparable. Although the function of the human *TYRP1* gene product remains undefined, it does play a role in pigmentation as mutations in the *TYRP1* gene are linked to the hypopigmentation observed in ROCA individuals.

1.3.2.1.2 The tyrosinase-related protein 2 locus

The human *TYRP2* gene has been cloned and sequenced [Bouchard *et al*, 1994; Yokoyama *et al*, 1994]. It is located on chromosome 13q31-32 [Sturm *et al*, 1994] and has 8 exons that encode a 519 amino acid protein [Sturm *et al*, 1995]. The nascent form of the polypeptide is approximately 65kDa in size, which is glycosylated to give an 80kDa final product [Hearing and Tsukamoto, 1991; King *et al*, 1995]. The human *TYRP2* has 80% DNA sequence identity with the mouse gene [Jackson *et al*, 1994b]. The mouse gene encoding TYRP2 maps to chromosome 14 [Jackson *et al*, 1992]. Both the human and mouse proteins have DOPAchrome tautomerase activity [Bouchard *et al*, 1994]. This enzyme is also expressed in the forebrain, implying that it may have other functions [King *et al*, 1995].

The *TYRP2* gene was the second melanogenic enzyme gene to be cloned. The *TYRP2* locus is homologous to the mouse *slaty* locus. The only known mutant of the *slaty* locus is the *slt* allele, which results in a reduction of DOPAchrome tautomerase activity. This then results in a reduction of coat colour [King *et al*, 1995]. The point mutation in the *slaty* mutant is in one of the copper binding sites of the gene [Jackson *et al*, 1992]. Mice homozygous for this mutation on a nonagouti background have a slightly diluted coat and yellowing ears. Two further slaty mutations have been defined: the *slaty-2J* (*slt2J*), which produces the same phenotype as the original *slt* mutant, and the *slaty light* (Slt^{lt}), which has a more severe effect and is semidominant [Bud and Jackson, 1995].

Studies in human melanocytes have shown that pigmented melanocyte cells have readily detectable levels of *TYRP1* mRNA [Del Marmol *et al*, 1993] and low levels of *TYRP2* mRNA

[Sturm *et al*, 1995]. They have also found that after UV-B exposure the levels of products of *TYR* and *TYRP1* genes increased while that of *TYRP2* decreased. However, Tobin and colleagues (1994) reported an increase in TYRP2 protein, as well as TYR and TYRP1, in response to UV-B exposure of human epidermal melanocytes. Abdel-Malek *et al* (1994) found that the stimulation of melanogenesis in human melanocytes by UV light involved an increase in tyrosinase activity but no change in the amounts of tyrosinase or TYRP1. Western blot analysis showed a decrease in TYRP2. Some investigators have found an increase in activity of DOPAchrome tautomerase in the presence of MSH, while others have not [King *et al*, 1995]. The exact mechanism involved still needs to be uncovered.

1.3.2.2 The *P* locus

The *P* gene has been cloned and mapped to chromosome 15q11.2-12 in humans [Lyon *et al*, 1992; Lee *et al*, 1995]. It is homologous to the mouse *pinkeyed-dilution* (*p*) locus (Table 1.3) on mouse chromosome 7 [Lyon *et al*, 1992]. The human *P* gene has 25 exons, the first of which is non-coding [Lee *et al*, 1995]. It encodes an 838 amino acid polypeptide with 12 transmembrane domains, and is similar in structure to ion or small organic molecule transporters [Lee *et al*, 1995; Puri and Brilliant, 1998; Sturm *et al*, 1998]. Mutations in the human *P* gene result in OCA2 [Lee *et al*, 1994]. Angelman syndrome and Prader-Willi syndrome are two hypopigmentation disorders that involve deletions that sometimes encompass the *P* gene [Rinchik *et al*, 1993; Lee *et al*, 1994].

Recessive alleles of the mouse *p* gene reduce pigmentation, mainly eumelanin. At least 12 alleles of the *p* locus have been described in mice and some mutations affect reproduction, development and behaviour. Some of the alleles that are known are: pink-eyed dilution (*p*), which results in light eyes and coat; pink-eyed dilution-J (p^J) with similar phenotype to *p/p*; pink-eyed unstable (p^{un}) that starts off as p^{un}/p^{un} and reverts to wild-type; dark pink-eye (p^d) can occur due to X-ray exposure and results in intermediate pigmentation; pink-eyed sterile (p^{bs}) which results in dark eyes, coat colour between p^d/p^d and *p/p*, and sterile males or poorly fertile females; and p-cleft palate (p^{cp}) which has similar colours as *p/p*, can be neonatal lethal, and results in poorly fertile females [Lyon *et al*, 1992].

The function of the P protein in the mouse and human was thought to be tyrosine transport into the melanosomes [Lee *et al*, 1995]. Lee *et al* (1995) suggested that the P polypeptide in humans could be part of a family of transporters that may also transport tyrosine. However, a study by Gahl *et al* (1995) showed that the mouse *p* gene product does not have a significant tyrosine transport function even though a transport system does exist in the melanosome membrane. The other functions that have been suggested are: for transport of substances between the melanocyte cytoplasm and the melanosome lumen and the maintenance of the melanosome structure, as some melanosomes without the protein have disorganised structures [Lamoreux *et al*, 1995]. Recently, a study by Puri and Brilliant (1998) demonstrated that the p protein regulates the pH within melanosomes. Melanocytes of wild-type mice had melanosomes with a low pH (acidic) and the melanosomes of p-deficient cells were almost never acidic. Tyrosinase activity is dependent on low pH and the higher pH in p-protein deficient cells may lead to minimal melanin production. This high pH may also have an adverse effect on the formation of the melanosomal complex [Puri and Brilliant, 1998].

Genetic and biochemical evidence suggests that TYR, TYRP1 and TYRP2 form a high molecular weight *melanogenic complex* within the melanosome [Barton *et al*, 1988; Lamoreux *et al*, 1995]. The P protein has been shown to be a part of this melanogenic complex, which is associated with the melanosomal membrane [reviewed in Sturm *et al*, 1998]. High-molecular-weight forms of the three enzymes are absent in eumelanic ocular tissue of p^{un}/p^{un} (pink-eyed unstable) mice that cannot produce the normal P mRNA, but are normal at the loci for the three enzymes [Lamoreux *et al*, 1995]. Thus, normal levels of wild-type P protein are necessary for eumelanogenesis and the absence of this protein can contribute to the switch to phaeomelanin production. Hence, the melanogenic complex may require the P protein to maintain its integrity. The *P* mRNA is absent in phaeomelanic tissue, suggesting that the switch from eumelanin to phaeomelanin production has something to do with the absence of the P protein in addition to the role that *MC1R* plays in this process [Lamoreux *et al*, 1995]. A study by Oskam *et al* (1995) revealed the presence of a protein (38L) in the leprosy causing *Mycobacterium leprae*, which has significant similarity to the mouse and human P protein. This may somehow be involved in the depigmentation process in leprosy. However, the role that the 38L protein plays in the organism may provide an avenue for studies on the function of the P protein.

1.4 Molecular basis of normal pigmentation

In order to study the molecular basis of normal pigmentation various populations need to be studied, with respect to genes involved in the pigmentation process. Thus, the mechanisms that take part in the pigmentation of skin, hair and eyes have to be compared between differently pigmented individuals. The tyrosinase activity, melanosome structure and content, skin typing, skin reflectance, climatic adaptations and other features that differ between different people have been discussed. The basic factor that brings about the phenotypic variation is the variation in the genes. The combined influences of the environmental factors and the genetic factors give rise to the pigmentary differences in normally pigmented individuals. Hence, it is essential to study the genetics of normal pigment variation. Previous studies on normally pigmented individuals will be discussed, as well as the conclusions that have been drawn and a proposal of further studies to clarify the complex mechanism of pigmentation.

1.4.1 The *MC1R* locus

The *MC1R* gene has been linked by many studies to the red hair and pale skin phenotype. This will be discussed along with the variation that has been observed at this locus in other normally pigmented individuals that do not have that particular phenotype.

1.4.1.1 The red hair and pale skin phenotype

Individuals with red hair have predominantly phaeomelanin in their skin and hair and/or reduced ability to produce eumelanin [Little and Wolf, 1981; Thody *et al*, 1991]. Red hair follicles have been found to exhibit phaeomelanogenesis or mixed type melanogenesis [Jimbow *et al*, 1983]. In the past many groups have tried to link red hair to various factors and loci. In 1935, Penrose found linkage of the phenotype to the ABO blood groups, which was later refuted [Penrose, 1935, and Hauge and Helweg-Larsen, 1954- cited in Eiberg and Mohr, 1987]. Red hair colour in 832 Danish families was also thought to be linked to the MNS blood groups, on chromosome 4 [Eiberg and Mohr, 1987]. More recently, however, many studies have shown strong linkage of this pigmentation type to the *MC1R* locus.

A study by Valverde *et al* (1995) revealed that changes in the sequence of the *MC1R* gene are linked to the red hair and pale skin phenotype. Approximately 80% of 30 red-haired individuals studied from Britain and Ireland had variation in this gene. These variants are listed in Table 1.4. Eight of them were in the second transmembrane domain and the ninth was in the seventh transmembrane domain [Valverde *et al*, 1995]. They listed the V92M as being associated with the phenotype, but it was later found in Chinese, Yakut and in Caucasoid individuals with various pigmentation types [Box *et al*, 1997; Rana *et al*, 1999], even though it was found to confer a decrease in MSH affinity [Xu *et al*, 1996]. In another recent study carried out in the USA individuals with red hair were found to have more *MC1R* variation than the Caucasoid individuals without red-haired studied [Rana *et al*, 1999]. These variants are also included in Table 1.4.

The study by Box *et al* (1997) confirmed Valverde *et al*'s (1995) findings that the *MC1R* locus variation has a close linkage to the red hair and pale skin phenotype. In this study both monozygotic and dizygotic twins in Queensland, Australia, were studied. Nine additional mutations were identified. Interestingly, this group also found that some dizygotic twins that were discordant for hair colour had the same *MC1R* genotypes. This suggests that variation in the *MC1R* locus is necessary but not sufficient for the red hair and pale skin phenotype [Box *et al*, 1997]. A recent study by Smith *et al* (1998) on a random Irish population showed that the majority of the individuals with fair skin also had variants at the *MC1R* locus (Table 1.4).

Functional studies carried out by Frandberg *et al* (1998) on the *MC1R* gene with the R151C variant, which had been found previously in a red-haired individual, demonstrated the functional importance of this variant. Receptor binding and cAMP assays were carried out on the cells transfected with mutant and wild-type genes. Although the mutant and wild-type receptor had identical binding potencies, the response to α-MSH binding by increasing intracellular cAMP was less than with the wild-type protein. Hence, even though this variant had normal binding ability it could not be stimulated to produce a normal functional response. The cAMP transduction pathway, therefore, does not function due to the inability of the receptor to couple to the GTP-binding protein. This variation occurs in the second intracellular loop, which inhibits the generation of second messenger cAMP, which in turn results in less tyrosinase and the

synthesis of phaeomelanin [Frandberg *et al*, 1998]. Another study investigated the effect of mutagenesis of *MC1R* on its binding potency to NDP-MSH (a superpotent α-MSH molecule) and found that certain residues of the protein were important for the binding of these two molecules [Yang *et al*, 1997]. Hence, mutations at these residues resulted in less binding potential between the hormone and receptor.

These studies confirm that variation at the *MC1R* locus plays some role in the red hair and pale skin phenotype. The functional study by Frandberg and colleagues provides evidence that mutations at the *MC1R* locus do affect MC1R protein function. The other variants observed might also bring about changes in the protein that make it unable to function normally. In South Africa a study on individuals with red-hair has not been carried out. South Africa consists of a mixture of individuals that originated from different parts of the world. Hence, it would be of interest to investigate the variation that occurs at the *MC1R* locus in South African Caucasoid individuals with red hair. Such a study may reveal other novel mutations that previous studies could not identify because their sample was restricted to one area.

1.4.1.2 Other types of normal pigmentation

Variants of the *MC1R* locus have also been found to occur in normally pigmented individuals. Valverde *et al* (1995) found variation in *MC1R* in individuals with other pigmentation types, including individuals with auburn, blond, brown, and black hair. However, the variants in *MC1R* were only found in the heterozygous state together with a normal allele. All the variants occurred at low frequencies. *MC1R* variants were correlated with skin types I and II, and only two individuals with skin type III and none with type IV had variants in this gene [Valverde *et al*, 1995]. Smith *et al*'s (1998) study observed *MC1R* variation in individuals with darker hair and skin in the Irish population (Table 1.5), but no particular mutation was associated with the darker phenotype. Ichii-Jones *et al* (1998) suggested that *MC1R* does not appear to play a significant role in skin type in terms of allelic variants. They carried out a study on Caucasian controls and melanoma cases, where they investigated the frequency of three alleles (V92M, D294H and D84E) in association with skin and hair colour and found no significant association with regard to skin type.

Koppula *et al* (1997) studied 60 individuals with skin types I to IV and found that two polymorphic alleles (Table 1.5) were linked to skin type I. Individuals that had these variants did not have red hair, they were blond with blue eyes. This links *MC1R* gene variation to pale skin, predisposes to skin cancer, and may explain the occurrence of the pale skin in red-haired individuals [Koppula *et al*, 1997], or that certain mutations give rise to just the pale skin, and others to red hair and pale skin. It is also possible that certain variants affect epidermal melanocytes only, and that others affect follicular melanocytes as well, resulting in red hair.

Rana *et al* (1999) studied individuals of different r origins: including East and Southeast Asians (Chinese, Cambodians, Japanese, Mongolians and Vietnamese), Africans, American Indians and Europeans. The mutations found in the red-haired individuals studied by this group were not found in the other populations. The only significant variant was the R163Q mutation in the East and Southeast Asians (Table 1.5), where it had a frequency of 0.70. This variant was also detected in Chinese individuals in Queensland, where it was found in 16 of 20 chromosomes (frequency of 0.8) [Box *et al*, 1997].

The Khoisan individuals have a more yellow-brown pigmented appearance of the skin, which may be due to greater phaeomelanin production. Skin reflectance tests show that they have the lightest skin of an unmixed indigenous group in southern Africa [Nurse *et al*, 1985]. Hence, *MC1R* variation may be a contributory factor to this phenotype, as mutations in the gene have been shown to contribute to yellow coat colour in various animals. A *MC1R* study in the San and Negroid may help us to understand the role of this gene in the normal pigmentation types of these groups.

1.4.2 The TYRP gene family and normal pigmentation

Studies have shown that mutations in the *TYR* gene result in OCA1. However, not many population studies have been carried out to assess the possible role that *TYR* gene variation has to play with regard to normal pigmentation. The polymorphisms that have been found in *TYR* are S192Y [Gieble and Spritz, 1990] and R402Q [Tripathi *et al*, 1991, cited in Spritz and Hearing 1995] and they are present in all populations with the exception of Orientals. African populations

Koppula *et al* (1997) studied 60 individuals with skin types I to IV and found that two polymorphic alleles (Table 1.5) were linked to skin type I. Individuals that had these variants did not have red hair, they were blond with blue eyes. This links *MC1R* gene variation to pale skin, predisposes to skin cancer, and may explain the occurrence of the pale skin in red-haired individuals [Koppula *et al*, 1997], or that certain mutations give rise to just the pale skin, and others to red hair and pale skin. It is also possible that certain variants affect epidermal melanocytes only, and that others affect follicular melanocytes as well, resulting in red hair.

Rana *et al* (1999) studied individuals of different racial origins: including East and Southeast Asians (Chinese, Cambodians, Japanese, Mongolians and Vietnamese), Africans, American Indians and Europeans. The mutations found in the red-haired individuals studied by this group were not found in the other populations. The only significant variant was the R163Q mutation in the East and Southeast Asians (Table 1.5), where it had a frequency of 0.70. This variant was also detected in Chinese individuals in Queensland, where it was found in 16 of 20 chromosomes (frequency of 0.8) [Box *et al*, 1997].

The Khoisan individuals have a more yellow-brown pigmented appearance of the skin, which may be due to greater phaeomelanin production. Skin reflectance tests show that they have the lightest skin of an unmixed indigenous group in southern Africa [Nurse *et al*, 1985]. Hence, *MC1R* variation may be a contributory factor to this phenotype, as mutations in the gene have been shown to contribute to yellow coat colour in various animals. A *MC1R* study in the San and Negroid may help us to understand the role of this gene in the normal pigmentation types of these groups.

1.4.2 The TYRP gene family and normal pigmentation

Studies have shown that mutations in the *TYR* gene result in OCA1. However, not many population studies have been carried out to assess the possible role that *TYR* gene variation has to play with regard to normal pigmentation. The polymorphisms that have been found in *TYR* are S192Y [Gieble and Spritz, 1990] and R402Q [Tripathi *et al*, 1991, cited in Spritz and Hearing 1995] and they are present in all populations with the exception of Orientals. African populations

have not been investigated and there is no data that suggests a connection between variation in this gene and normal variation in pigmentation. A study on *TYR* gene variation in a group of Negroid and Caucasoid individuals, the two extremes of the normal pigment spectrum, may help to determine the role of this gene in normal pigmentation.

Box *et al* (1998) carried out a study to attempt to link normal pigment variation to the *TYRP1* locus by investigating 100 normally pigmented Caucasoid individuals. The only polymorphism they detected in the coding region was a synonymous base substitution in 11 of 200 alleles at R87 in exon 2 of the *TYRP1* gene [Box *et al*, 1998]. No other studies, to my knowledge, have been carried out on this gene in normally pigmented individuals. A study of Negroid and Caucasoid individuals may shed more light on the involvement of this gene in pigment variation. The *TYRP2* gene was studied by Box and colleagues and was not shown to play a significant role in normal pigmentation [personal communication]. As with the *TYR* gene, a comparison of variation in the *TYRP1* and *TYRP2* genes in the two extremes of human pigmentation, the very dark Negroid and very light Caucasoid populations, may reveal more on the mechanism of pigmentation at a molecular level.

Other loci are likely to be involved in this polygenic trait of pigmentation. Studies have been carried out to investigate brown eye colour (BEY1) and it has been linked, by linkage studies, to blood groups Colton (CO) and Kidd (JK) [Gedde-Dahl *et al*, 1982, cited in Eiberg and Mohr, 1996]. The phenotype of green eye colour (GEY) has been mapped to chromosome 19 [Eiberg and Mohr, 1996]. The *BEY2* gene has, by linkage analysis, been assigned to region 15q11-21, a region close to the *P* gene. A locus for brown hair colour (*HCL3*) was also localised because of its linkage to *BEY2* [Eiberg and Mohr, 1996]. Hence, even though *α-MSH*, *MC1R*, *TYR*, *TYRP1*, *TYRP2* and the *P* genes may play a significant role in pigmentation other loci may exist that play a part in this complex trait.

Table 1.4: *MC1R* variation observed in Caucasoid individuals with red hair by various studies

Variant	Position in Protein	Origin of individuals	Other	References
A64S	First intracellular loop	Britain and Ireland		Valverde *et al* (1995)
K65N	First intracellular loop	Queensland, Australia		Box *et al* (1997)
F76Y	Second transmembrane domain	Britain and Ireland		Valverde *et al* (1995)
T95M	Second transmembrane domain	Britain and Ireland		Valverde *et al* (1995)
V97I	Second transmembrane domain	Britain and Ireland		Valverde *et al* (1995)
A103V	First extracellular loop	Britain and Ireland		Valverde *et al* (1995)
L106Q	First extracellular loop	Britain and Ireland		Valverde *et al* (1995)
R142H	Second intracellular loop	Queensland, Australia		Box *et al* (1997)
R151C	Second intracellular loop	Queensland, Australia Ireland USA		Box *et al* (1997) Smith *et al* (1998) Rana *et al* (1999)
R160W	Fourth transmembrane domain	Queensland, Australia Ireland		Box *et al* (1997) Smith *et al* (1998)
D294H	Seventh transmembrane domain	Britain and Ireland Queensland, Australia Ireland	Loss of *TaqI* site	Valverde *et al* (1995) Box *et al* (1997) Smith *et al* (1998)
A299T	Seventh transmembrane domain	Queensland, Australia		Box *et al* (1997)

Table 1.5: *MC1R* gene variation observed in other types of normal pigmentation

Variant	Position in Protein	Population group	Other	References
V60L	First transmembrane domain	Caucasoids with various hair colours		Box *et al* (1997) Smith *et al* (1998)
R67V	First intracellular loop	Chinese East and Southeast Asian		Box *et al* (1997) Rana *et al* (1999)
D84E	Second transmembrane domain	Caucasoid with any colour hair	Loss of *AvaII*	Box *et al* (1997) Koppula *et al* (1997) Smith *et al* (1998) Rana *et al* (1999)
V92M	Second transmembrane domain	Caucasoids with various hair colours	Gain of *NlaIII* Gain of *NspI*	Valverde *et al* (1995) Box *et al* (1997) Koppula *et al* (1997) Rana *et al* (1999)
V92L	Second transmembrane domain	Caucasoid with any colour hair		Box *et al* (1997)
I155T	Second intracellular loop	Caucasoids with dark hair Caucasoids with red hair		Smith *et al* (1998) Box *et al* (1997)
R163Q	Second intracellular loop	Chinese East and Southeast Asian Caucasoids with fair or dark hair		Box *et al* (1997) Rana *et al* (1999) Smith *et al* (1998)
T314T	COOH terminal region	Caucasoid(s) African East and Southeast Asians Indians		Box *et al* (1997) Rana *et al* (1999)

1.5 Aims

1. To screen for DNA sequence variation in the *MC1R* gene and to identify mutations associated with pigmentation differences in a random group of Negroid and San individuals.
2. To confirm the role of the *MC1R* gene in Caucasoid individuals with red hair and pale skin by investigating a random group of South African red-haired individuals of European origin.
3. To screen the five exons of the *TYR* gene by SSCP analysis and sequencing in a group of random normally pigmented Negroid and Caucasoid individuals in an attempt to find alleles with possible functional significance that may play a role in normal pigment variation.
4. To investigate the role of the *TYRP1* gene in normal pigmentation by:
   a) investigating the frequency of two Rufous OCA mutations (S166X and 368delA) in random Negroid, Caucasoid, San and red-haired individuals,
   b) screening the entire gene for significant variation, by SSCP analysis and sequencing, between groups of Negroid and Caucasoid individuals in order to assess its role in normal pigment variation.

2. SUBJECTS AND METHODS

2.1 Subjects

DNA samples from random Negroid, Caucasoid and San individuals were studied, as well as a selected group of Caucasoid individuals with red hair. Most of the samples were obtained from the DNA repository available at the department, which has been set up as part of a large study on the genetic diversity of sub-Saharan populations. These samples were obtained after ethics approval and verbal consent was obtained from all individuals. Normal variation at the *MC1R* locus was investigated by studying a group of random Negroid, San and Caucasoid individuals with red hair. *TYR* and *TYRP1* genes were studied by investigating a random group of normally pigmented Negroids and Caucasoids to investigate normal DNA variation at these loci.

2.1.1 Random Negroid groups

A large number of the Negroid DNA samples were obtained from the repository, but on occasion blood samples were processed to obtain the DNA. The Negroid samples were obtained from two different sources. One was used for sequence variation detection at the *MC1R* locus and for the ROCA mutation studies, both of which were carried out at the start of this project. The second random Negroid group was obtained to study differences at the *TYR* and *TYRP1* loci in normal pigmentation, as a part of a new project in the department that commenced towards the middle of the second year.

Paternity samples used for MC1R studies

Tests on paternity determination provided a useful resource of samples from random individuals belonging to different population groups. A group of 25 nuclear families from the Negroid population, including the mother, father and at least one child were identified. The aim was to study the mother and father, who are only related by marriage, to find normal variation in the *MC1R* gene. If variation was found such that the parents were doubly heterozygous for mutations at the locus the child would be tested to determine the phase (cis- or trans-) of the variants. Unfortunately, due to a misunderstanding on sample labelling and

allocation, the children's samples were thought to be the fathers' samples and this error was only detected after most of the project was done. Due to limited recourses the typing for *MC1R* variation could not be carried out in the fathers and the results were, therefore, analysed to reflect random chromosomes, i.e. three per family (two from the child and one that was not inherited by the child from the mother). All 50 normally pigmented Negroid parent samples were, however, used to study the frequency of the 368delA and S166X ROCA mutations, at the *TYRP1* locus, in normally pigmented Negroid individuals.

Random Negroid samples used for TYR and TYRP1 screening

Fresh blood samples were obtained from 15 normally pigmented Negroid individuals at the Baragwanath hospital in Johannesburg, and the DNA was extracted. *TYR* and *TYRP1* studies were carried out to examine normal sequence variation in these Negroid individuals.

Random San individuals

A group of 17 random San (!Kung) DNA samples from the DNA repository were studied. These samples were originally collected from Tsumkwe, Namibia, in 1985. The San have a natural yellow tinge in their skin, which may suggest *MC1R* or *TYRP1* mutations. They were, therefore, screened for *MC1R* sequence variation and for the ROCA mutations at the *TYRP1* locus.

2.1.2 Random Caucasoid groups

A total of three groups of Caucasoid individuals were studied. Two groups were random normally pigmented Caucasoid samples: one group was used for the ROCA mutation study and the second was used to screen the *TYR* and *TYRP1* genes for normal sequence variation. The third group was comprised of Caucasoid individuals with red hair who were used to study *MC1R* gene variation.

The first was a group of 50 unrelated South African Caucasoid individuals, European descent, from the departmental collection that was used for the ROCA mutation study. The second group included 15 random Caucasoid individuals from the Department of Human Genetics, at the SAIMR, and were used for *TYR* and *TYRP1* gene studies on normal pigment

variation. The latter samples were obtained as fresh blood samples, from which the DNA was extracted.

2.1.3 Caucasoid group of individuals with red hair

Caucasoid individuals with red hair were from the Johannesburg area. Ethics clearance was obtained from the committee for research in human subjects, protocol number M 970510 (Appendix 1), before individuals were approached and provided with the information sheet explaining the aims of the project. They were given a consent form to sign and a qualified nurse or doctor obtained a sample of blood from them. Samples were obtained from seven unrelated red-haired individuals and one of their mothers. They were screened for mutations in the *MC1R* gene and the ROCA mutations at the *TYRP1* locus.

2.2 Methods

Most of the techniques used in this project are related to mutation detection. Where DNA was not already available, it was extracted from the white blood cells by the salting-out procedure. The sequences of interest were amplified by the polymerase chain reaction (PCR), the product of which was purified when necessary and subjected to the various mutation detection procedures to identify DNA sequence variation. The recipes for the solutions used in this project are given in Appendix 2, followed by a list of the reagents and their suppliers in Appendix 3.

2.2.1 DNA extraction by the salting out procedure

Miller *et al* (1988) described the salting out procedure for extracting DNA from white blood cells. The blood samples were collected in tubes containing EDTA or ACD as anti-coagulants. The blood was decanted into polypropanol tubes and stored at –20°C. Alternatively, the samples were centrifuged at 2 300rpm in a bench top centrifuge (Beckman) at 4°C for 10 minutes. The layer of white blood cells (the 'buffy layer') was aspirated from the centrifuged samples and used for the extraction of DNA or stored long term at –20°C.

DNA extraction from frozen whole blood or buffy samples is as follows. The tubes containing samples were thawed and filled to the 45-50ml mark with cold sucrose-triton-X lysing buffer (Appendix 2). This buffer swells the red blood cells causing the membranes to rupture. The tubes were inverted several times to mix the solutions and centrifuged at 2 300rpm for 10 minutes at 4°C. The supernatant fluid containing residual red cell debris was discarded. The pellet, containing white blood cells, and possibly more red blood cells, was washed again in 20-25ml sucrose-triton-X lysing buffer and placed on ice for 5 minutes. The tubes were then centrifuged for 5 minutes at 2 300rpm, at 4°C, after which the supernatant fluid was discarded. To the pellet 3ml T20E5, 0.2ml 10%SDS and 0.5ml Proteinase-K mix were added (Appendix 2) and mixed by inversion. This was incubated at 42°C overnight. The SDS is a detergent that breaks down the lipid layer of the cell membrane. The EDTA in the T20E5 solution chelates and removes Mg^{2+} ions, which are necessary for the function of enzymes such as RNases and DNases. The proteinase K enzyme breaks down the protein of the cell at an optimal temperature of 42^0C.

One ml saturated NaCl (Appendix 2) was added to the mix and agitated by inversion, followed by incubation on ice for 5 minutes. The tubes were centrifuged at 4°C at 2 300rpm, for 30 minutes. At this point the salt binds to the protein and precipitates to form a white pellet. The supernatant contained the DNA, which was decanted into a new tube. Approximately two volumes of absolute ethanol (100%) were added to the supernatant. The solution was agitated to precipitate the DNA. This DNA was spooled, 'fished out' and washed in cold 70% ethanol. The ice-cold ethanol removes the NaCl bound to the DNA. The DNA was then air dried and dissolved in an appropriate volume of 1xTE (Appendix 2). In cases where the spooling procedure did not produce a good yield, the DNA from the sample was precipitated by over night storage at -20°C. After centrifugation for half an hour at 2 300rpm (4°C), the supernatant was discarded and the pellet was diluted in 1xTE. The quality of the DNA samples was checked by agarose gel electrophoresis. If the DNA were degraded a smear was seen on the gel. If, as in most cases, the DNA was not degraded a strong high molecular weight band was observed close to the wells of the gel.

2.2.2 The polymerase chain reaction (PCR)

PCR is an *in vitro* procedure by which a specific DNA sequence is amplified by using primers that flank the region of interest that are complementary to the two strands of DNA. Mullis and Faloona first described PCR in 1987. It is especially useful when small amounts of DNA are present for analysis. PCR is used as a first step to many molecular techniques, and was necessary in all studies carried out here with the *MC1R*, *TYR* and *TYRP1* genes.

Two primers are designed, each being complementary to each of the two strands of the double stranded DNA and flanking the sequence of interest. The following reagents are necessary for PCR in addition to the template DNA and primers: deoxyribonucleotides (dNTPs); DNA polymerase, the enzyme that synthesises the complementary DNA strand in the 5' to 3' direction; and a buffer [Taylor, 1992]. Taq polymerase is the enzyme that is used because it is thermostable and can withstand the high temperatures that are necessary for PCR. The buffer is commercially manufactured and contains Tris-HCl, $MgCl_2$, KCl, and gelatine, all of which are essential for PCR as they help with primer annealing and enzyme activity [Innis and Gefald, 1990]. Additional reagents such as spermedine, dimethyl sulfoxide (DMSO), formamide and/or varying $MgCl_2$ concentration have been found to be necessary for certain PCR systems. The additional reagents work individually or together to improve the yield of the PCR product.

PCR involves alternative heating and cooling cycles. First, the double stranded DNA is denatured at 90-95°C to obtain single stranded molecules, making the bases more accessible by the primers. The temperature is then dropped to a suitable *annealing* temperature at which the primer can anneal to the DNA strand. This depends on the primer sequence, which dictates its melting temperature [a rough estimate of the annealing temperature is calculated by T_m= ((A+T)2+(C+G)4)-5)]. The chosen annealing temperature has to be below the primer's melting temperature. Finally, the cycle goes through the *extension step*, during which the polymerase extends the primer with dNTPs to produce a DNA strand complimentary to the template DNA strand. These three basic PCR steps are repeated between 30 and 40 times so as to obtain a large amount of the sequence of interest.

MC1R has only one exon, and therefore a single PCR reaction was sufficient to amplify the entire gene. *TYR* and *TYRP1* genes are, on the other hand, composed of 5 and 8 exons respectively. Each of these exons was amplified individually from genomic DNA. In fact, exon 1 of *TYR* is large and was amplified in two separate PCR reactions (as exons 1A and 1B). The optimised conditions for each of these are presented in Table 2.1. The recipes of the solutions that were used in this procedure are presented in Appendix 2.

DNA (1-2µl containing approximately 50-100ng of DNA) was aliquoted into appropriately labelled PCR tubes. A suitable volume of reaction mix, containing the necessary PCR reagents, was then added to the DNA. The reaction mixes contained the following reagents: 0.25-1µl primers (10pmol/µl); 2.5µl dNTP (125µM) with dATP, dCTP, dGTP and dTTP in equal proportions; 2.5µl PCR buffer (with the 10x stock containing 100mM Tris-HCl, 15mM $MgCl_2$, 500mM KCl, 1mg/ml gelatine; pH 8.3 at 20°C); 0.2µl Taq polymerase enzyme (5U/µl) and double-distilled water for a 25µl reaction. The reaction mix for exon 6 of the *TYRP1* gene also included 0.5µl spermedine (2mM) in the 25µl reaction. The DNA and primer concentrations varied between most systems. The primer concentrations are presented in Table 2.1. Once the reaction mix was added to the DNA sample, oil was added to prevent the evaporation of the PCR reaction mix, and the tubes were placed in an automated thermocycler (Cetus, Hybaid Ominigene or Hybaid Touchdown) and subjected to the appropriate optimised PCR conditions (Table 2.1).

A radioactively labelled dNTP [12.5µCi of ^{32}P-dCTP (250µCi/µl)] was added to the reaction mix (total volume of 25µl) when the PCR product was to be used for *single stranded conformation polymorphism* (SSCP) analysis, so as to detect conformational changes in DNA sequences. When working with radioactive substances it was vital to work carefully behind Perspex screens. This procedure was used for *TYR* and *TYRP1*, when screening the exons for DNA sequence variation.

Table 2.1: List of PCR primers and reaction conditions

Gene	Exons	Primer sequence (5' to 3')	Product (bp)	Primer conc. (pmol/µl) *	PCR conditions (thermocycler)	Reference
MC1R		1F -AGATGGAAG AGGCAGGCAT 1R -CCGCGCTTCAACACTTTCAGAGATCA	1238	0.1	95°C 1', 55°C 1', 72°C 1' for 40 cycles (Cetus)	Box *et al*, 1997
TYR	1A	1AF -TGACTCCAATTAGCCAGTTC 1AR -ACATAGTCTGAGCTGATGC TA	499	0.2	94°C 1', 60°C 1', 72°C 1' for 30 cycles, 72°C 10' (Omnigene) This was the same for all exons except 2 and 3, which had an annealing temperature of 55°C	
	1B	1BF -TTTGCCTACCTCACTTTAGC 1BR -TTAACAGGGCACCATTTCTG	548	0.2		
	2	2F -CTCAGGAGAAGTCTAACAAC 2R -AACTCAGAAATTCTGAATTC	420	0.4		
	3	3F -AGTCTCAATACGGAATGAA 3R -TTTAAATCCAATGAGCACGT	336	0.4		
	4	4F -AACATCTTTCCATTGTCTCCAG 4R -CACTAGATTCAGCAATTCCTCT	338	0.2		
	5	5F -TTAGGTGTAACTTTCCCAAGC 5R -GAAGGCTACACTTTGTATTAGG	522	0.2		
TYRP1	1	- (non-coding)			94°C 1', annealing 1', 72°C 1' for 30 cycles and 72°C 10'	Box *et al*, 1998
	2	2F -CGTGCTTCAGTCTTCTCTACA 2R -GCAGGACTTATGAACTCATTC	490	0.4	55°C (Cetus)	
	3	3F -CGCAAGGCAGATGTTTTCATG 3R -AAGGCATCTTGTCTGTAAAGA	416	0.4	57°C (Cetus)	
	4	4F -AGACCAAACGAAAATGAATA 4R -AAATTCTGACTCCAAGCTATC	306	0.2	47°C (Omnigene)	
	5	5F -AAAGAGCGACAATAAGAACTC 5R -AAAGCCTTCTCAAAGAAACTT	319	0.2	50°C (Hybaid)	
	6	6F -TTGCTATTACCTGGAAAAGTG 6R -TGCAAAAAGCATATGAAAATG	275	0.04	51°C (Cetus)	
	7	7F -ATACGTTGTCTTTGGAATAAT 7R -ATACCGTGATTACTCTACTTG	252	0.2	45°C (Hybaid)	
	8	8F -TGTCCACTTTTTGGTGATAAC 8R -ATTCAACCAGGTGGTTTTGTG	324	0.4	51°C (Cetus)	

* The primer concentration refers to the concentration of the primer in 25µl reaction.

2.2.3 Restriction endonuclease digestion

Restriction endonucleases can be defined as enzymes that recognise specific nucleotide sequences in a DNA molecule, and cleave at that site. Type II restriction enzymes recognise palindromic sequences (those that read the same in both directions) and cleave to produce double stranded breaks. The site at which the restriction enzyme cleaves is called a *restriction site*. The cleaving produces *restriction fragments* of known sizes. Restriction endonuclease digestion was used for the detection of the S166X ROCA mutation and in the *TYR* and *TYRP1* studies, in order to obtain DNA fragments of suitable size for SSCP analysis. The restriction enzymes used in this project are listed in Table 2.2.

Restriction enzyme digestion is the basic procedure for *restriction fragment length polymorphism* (RFLP) analysis and was used to detect the S166X ROCA mutation in the *TYRP1* gene. The procedure will be described in more detail in section 2.2.7.2. This method of enzymatic cleavage was also necessary for SSCP analysis of large fragments. Variations in the DNA sequence can only be detected optimally in fragments of approximately 300bp or smaller in size. Hence, the PCR products much larger than 300bp were cleaved by appropriate restriction enzymes. The restriction enzymes used in each case are listed in Table 2.2. The process of SSCP will be discussed in detail in section 2.2.7.1. The *TYR* products that had to be cleaved were from exons 1A (498bp), 1B (548bp), 2 (420bp) and 5 (522bp). The *TYRP1* exons whose products were cleaved were from exons 2 (490bp) and 3 (416bp). Table 2.3 shows the product sizes obtained for each system after restriction enzyme digestion.

In the *TYR* and *TYRP1* systems, 9µl of the PCR product was digested with 1µl enzyme (10U) and 1.5µl of the buffer in a volume of 15µl. The buffer is specific for each enzyme, and contains all the appropriate reagents for the normal and optimum functioning of the enzyme. The samples were incubated at the appropriate temperature for two hours in a Maxi 14 hybridisation oven (Hybaid). Different enzymes act optimally at different temperatures, but, fortunately, all the enzymes used in this study worked optimally at 37°C.

Table 2.2: Restriction enzymes used and their recognition sites

Restriction enzyme	Recognition site	Buffer used and contents	Use in study
Rsa I	GT|AC	(L) 10mM tris-HCl, 10mM $MgCl_2$, 1mM dithioerythritol	SSCP exon 1B of *TYR* with *Hinf I*
Hinf I	G|ANTC	(H) 50mM tris-HCl, 10mM $MgCl_2$, 100mM NaCl, 1mM dithioerythritol	SSCP exon 1B of *TYR* with *Rsa I* SSCP exon 2 of *TYR* with *Ava II* SSCP exon 5 of *TYR* with *Bgl I* SSCP exon 2 *TYRP1* with *Dde I*
Bgl I	GCCN|NGGC	(M) 10mM Tris-HCl, 10mM $MgCl_2$, 50mM NaCl, 1mM dithioerythritol	SSCP exon 2 of *TYR* with *Hinf I* SSCP exon 5 of *TYR* with *Hinf I*
Ava II	G|GA/TCC	(A) 33mM tris-Ac, 10mM MgAc, 66mM KAc, 0.5mM Dithiothreitol	SSCP exon 1A of *TYR* with *Hae III*
Hae III	GG|CC	(L) as for *Rsa I*	SSCP exon 1A of *TYR* with *Ava II*
*MboI**	|GATC	(K)* 80mM salt, 30% NaCl, 40% KCl	Detection of S166X mutation
Dde I	C|TNAG	(M) as for *Bgl I*	SSCP exon 2 of *TYRP1* with *Hinf I* SSCP exon 3 of *TYRP1* with *Alu I*
Alu I	AG|CT	(A) as for *Ava II*	SSCP exon 3 of *TYRP1* with *Dde I*

*All enzymes and buffers were obtained from Boehringer Mannheim except for *MboI*, which was from Amersham

2.2.4 Agarose gel electrophoresis (AGE)

AGE is used for quantitative or qualitative purposes and was used in this study to check the integrity of the DNA samples used, to identify products of the different PCR systems and to identify the restriction fragments in the S166X investigation.

The technique involves the electrophoresis of the negatively charged DNA sample through a gel during which the DNA of different sizes moves through the gel. The distance that the DNA molecule moves in the gel depends on its size. First, the DNA is mixed with ficoll dye (Appendix 2), which causes the sample to be heavy and sink into the wells of the gel. An electric current is then applied to the gel so that the positive electrode (anode) is connected to the end and the negative electrode (cathode) is connected to the beginning of the gel (at the wells) [Davis, 1991]. DNA has a negative charge and, as a result, moves towards the positive electrode with the current. As the sample passes through the gel the larger pieces move with less ease through the pores of the gel than the smaller pieces. Therefore, the different sizes of molecules migrate to

different positions in the gel, with the larger fragments retarded closer to the wells and the smaller fragments further from the wells.

In this project gels ranging from 0.8% to 3% were used, the former for large products of extracted high molecular weight DNA and the latter for restriction enzyme digested products as small as 155bp. Th recipe for the preparation of the gel is described in Appendix 2. The DNA, PCR (2-5µl) or restriction digestion product (30µl) was mixed with 5µl ficoll dye and pipetted into the wells of the gel, which was placed in 1XTBE running buffer (Appendix 2). A voltage of 100V to 120V was applied in most cases for electrophoresis. The ethidium bromide in the gel intercalates with the DNA, which makes it visible under UV light. Therefore, after electrophoresis, the DNA in the gel was visualised under UV light on an UV-transilluminator. In the S166X study this process was crucial to observe the different restriction fragments to assess the genotypes of the individuals studied.

2.2.5 PCR product purification

In order to carry out sequencing it is necessary to purify the PCR product to ensure good template quality. As sequencing was carried out in the three genes investigated, this section applies to all three systems. The purification procedure removes excess dNTPs and primers used for PCR. Three techniques were used to purify PCR product: two involved excision of the correct product from a low melting temperature gel, while the third involved purification with enzymes.

2.2.5.1 Purification by band excision

The amplified product of certain PCR systems result in spurious bands, and the PCR product of interest is separated from the spurious products by being excised from a gel that the sample has been electrophoresed on. This procedure was necessary mainly for amplified *MC1R* and exons 6 and 7 of the *TYRP1* gene.

The PCR product is first electrophoresed through a low melting temperature gel and excised from it. The gel slice can then be purified with the Wizard purification kit (Promega) or

different positions in the gel, with the larger fragments retarded closer to the wells and the smaller fragments further from the wells.

In this project gels ranging from 0.8% to 3% were used, the former for large products of extracted high molecular weight DNA and the latter for restriction enzyme digested products as small as 155bp. Th recipe for the preparation of the gel is described in Appendix 2. The DNA, PCR (2-5µl) or restriction digestion product (30µl) was mixed with 5µl ficoll dye and pipetted into the wells of the gel, which was placed in 1XTBE running buffer (Appendix 2). A voltage of 100V to 120V was applied in most cases for electrophoresis. The ethidium bromide in the gel intercalates with the DNA, which makes it visible under UV light. Therefore, after electrophoresis, the DNA in the gel was visualised under UV light on an UV-transilluminator. In the S166X study this process was crucial to observe the different restriction fragments to assess the genotypes of the individuals studied.

2.2.5 PCR product purification

In order to carry out sequencing it is necessary to purify the PCR product to ensure good template quality. As sequencing was carried out in the three genes investigated, this section applies to all three systems. The purification procedure removes excess dNTPs and primers used for PCR. Three techniques were used to purify PCR product: two involved excision of the correct product from a low melting temperature gel, while the third involved purification with enzymes.

2.2.5.1 Purification by band excision

The amplified product of certain PCR systems result in spurious bands, and the PCR product of interest is separated from the spurious products by being excised from a gel that the sample has been electrophoresed on. This procedure was necessary mainly for amplified *MC1R* and exons 6 and 7 of the *TYRP1* gene.

The PCR product is first electrophoresed through a low melting temperature gel and excised from it. The gel slice can then be purified with the Wizard purification kit (Promega) or

QIAquick gel extraction kit (Quiagen). Both techniques were found to be equally efficient. For both procedures the whole volume of PCR product (25µl or 50µl) was electrophoresed on a Nusieve gel (1-3%) at 120mA from between 1 and 2 hours. The Nuseive gel has a low melting temperature and is made in 1XTAE (Appendix 2) or 1XTBE buffer. TAE is preferred as it gives better resolution of the band of interest. The bands were then cut out of the gel under UV light and placed in Eppendorf tubes to be purified.

2.2.5.1.1 Wizard purification

The gel slices containing DNA were heated at 70°C for 10 minutes. One ml of the Wizard mini-prep resin was added to the molten gel and mixed, during which the DNA binds to the resin. This mix was then added to a Wizard column that was fitted on a vacuum manifold. Applying the vacuum caused the molten gel to be sucked out with the DNA remaining bound to the resin, which in turn was bound to the column. The columns were washed twice with 1ml 80% isopropanol to ensure that all the gel was removed. They were then placed in collection tubes and centrifuged at 13 000rpm for 2 minutes to remove excess isopropanol. The columns were then placed in clean, labelled 1.5ml Eppendorf tubes and 50µl of double-distilled water was added to each one. This was left for 1 minute to allow the DNA to dissolve in the water, followed by centrifugation for 20 seconds at 13 000rpm. The elution contained pure PCR product dissolved in water. Five µl of this DNA was electrophoresed on an agarose gel (1-3%) in 1xTBE to determine its quality and quantity.

2.2.5.1.2 QIAquick gel extraction kit

The gel slices containing DNA were weighed and one volume of buffer QG was added to it. This was incubated at 50°C for 10 minutes to melt the gel. QG buffer contains a pH indicator, which allows for the easy determination of pH optimal for DNA binding. The pH of the solution is optimal at 7.5 or less, during which the colour of the solution is yellow. The mixture was vortexed every 2 minutes and 10µl 3M sodium acetate (Appendix 2) was added to each sample to improve the yield. One gel volume of 100% isopropanol was added to the latter mix if the solution was violet (pH >7.5) to bring it to the appropriate pH. This mixture was then applied to the QIAquick spin column containing a silica-gel membrane and centrifuged for 1 minute at 13 000 rpm. The DNA bound to the column while the contaminants passed through. The flow-

through was discarded and 500µl QG buffer was added to the column. This was centrifuged at 13 000rpm for 1 minute to remove all traces of gel. Then, 750µl PE buffer containing ethanol was added to the column and centrifuged for 1 minute at 13 000rpm after standing for between 2 to 5 minutes, to wash away salt. The flow-through was discarded and the columns were centrifuged again to ensure the removal of all traces of ethanol, which could interfere with subsequent enzymatic reactions. The purified product was then eluted with 50µl water. The product was electrophoresed on a 1-3 % agarose gel to determine its quality.

2.2.5.2 Shrimp alkaline phosphatase (SAP)-exonuclease purification

The enzyme method of PCR product purification is ideal for cases where the PCR does not produce spurious bands. This procedure was used in the *TYR* and *TYRP1* gene studies for most of the exons. The PCR product is purified of excess dNTPs and primer by this technique, as these may interfere with the sequencing process. Shrimp alkaline phosphatase (SAP) and exonuclease 1 enzymes remove these materials. Both are active in the PCR buffer and are, thus, added directly to the PCR products. Exonuclease 1 removes residual primer and any extraneous single-stranded DNA. The SAP removes the remaining dNTPs [PCR product pre-sequencing kit].

The appropriate amount of template for sequencing was aliquoted into a PCR tubes, which was "eyeballed" after electrophoresis on an agarose gel. Five µl of PCR product (100-200ng) was used in cases where a reasonably bright band was observed after the agarose gel electrophoresis. 0.5U of SAP and 0.5U of exonuclease were added to the PCR template, which was overlaid with oil. The samples were then subjected to incubation at 37°C for 15 minutes, during which the enzymes digested excess PCR components; and then 80°C for 15 minutes, during which the enzymes were inactivated. The temperature was regulated in the Hybaid Omnigene or Touchdown thermocyclers.

2.2.6 The dot blotting procedure

Dot blots were used for *allele specific oligonucleotide* (ASO) analysis, for the detection of the 368delA ROCA mutation in exon 6 of the *TYRP1* gene. The dot blotting procedure used is similar to that described by Theopilus *et al* (1989).

An appropriate amount of amplified PCR product of the gene of interest is first denatured and dotted on to a positively charged nylon membrane using a dot blotting apparatus. The negatively charged DNA binds to the membrane, which is dried and baked in an oven to cross-link the DNA permanently to the membrane. This blot is then used for hybridisation and detection.

The DNA samples were denatured in the following way. Ten µl of concentrated PCR product was aliquoted into Eppendorf tubes and 200µl denaturing solution (Appendix 2) was added and mixed. The samples were incubated at 65°C for 30 minutes. They were then neutralised by the addition of 237µl neutralising solution (Appendix 2). The optimum volumes for the denaturing and neutralising solutions were obtained by dotting varying volumes of the two solutions onto pH paper to obtain a neutral (pH7.0) mix. The positively charged nylon membrane ($88cm^2$) (Boehringer Mannheim) was then made ready by soaking in distilled water for 1minute and then in 2xSSC for 30 minutes before use.

The wet membrane was assembled onto the Hybri.Dot^{TM} Manifold (Bethesda Research Laboratories), which was connected to a vacuum pump. 2xSSC (250µl) was first pumped through the wells to be used (6mm in diameter) to ensure that the vacuum was working efficiently. Then, 500µl of the denatured samples were pipetted into the appropriate wells and vacuumed through. These wells narrowed to filter spots of 3mm in diameter, which was the final size of the dots on the membrane. The wells were then washed with 250µl 2xSSC to ensure that none of the DNA was left on the walls of the wells. The apparatus was disassembled and the blot was left to dry for at least 1.5 hours on Whattman paper (Merck) and baked at 120°C for 30 minutes in a vacuum oven (Gallenkamp) to link the DNA to the blot permanently. The blot was used for ASO analysis (section 2.7.2.3).

2.2.7 Mutation detection

The techniques that have been described thus far were prerequisites for the mutation detection procedures. Mutations such as single base changes, deletions and insertions can be detected using a variety of molecular techniques. These include denaturing gradient gel electrophoresis (DGGE), heteroduplex analysis, SSCP analysis and sequencing to identify unknown variation at

An appropriate amount of amplified PCR product of the gene of interest is first denatured and dotted on to a positively charged nylon membrane using a dot blotting apparatus. The negatively charged DNA binds to the membrane, which is dried and baked in an oven to cross-link the DNA permanently to the membrane. This blot is then used for hybridisation and detection.

The DNA samples were denatured in the following way. Ten µl of concentrated PCR product was aliquoted into Eppendorf tubes and 200µl denaturing solution (Appendix 2) was added and mixed. The samples were incubated at 65°C for 30 minutes. They were then neutralised by the addition of 237µl neutralising solution (Appendix 2). The optimum volumes for the denaturing and neutralising solutions were obtained by dotting varying volumes of the two solutions onto pH paper to obtain a neutral (pH7.0) mix. The positively charged nylon membrane ($88cm^2$) (Boehringer Mannheim) was then made ready by soaking in distilled water for 1minute and then in 2xSSC for 30 minutes before use.

The wet membrane was assembled onto the Hybri.Dot™ Manifold (Bethesda Research Laboratories), which was connected to a vacuum pump. 2xSSC (250µl) was first pumped through the wells to be used (6mm in diameter) to ensure that the vacuum was working efficiently. Then, 500µl of the denatured samples were pipetted into the appropriate wells and vacuumed through. These wells narrowed to filter spots of 3mm in diameter, which was the final size of the dots on the membrane. The wells were then washed with 250µl 2xSSC to ensure that none of the DNA was left on the walls of the wells. The apparatus was disassembled and the blot was left to dry for at least 1.5 hours on Whattman paper (Merck) and baked at 120°C for 30 minutes in a vacuum oven (Gallenkamp) to link the DNA to the blot permanently. The blot was used for ASO analysis (section 2.7.2.3).

2.2.7 Mutation detection

The techniques that have been described thus far were prerequisites for the mutation detection procedures. Mutations such as single base changes, deletions and insertions can be detected using a variety of molecular techniques. These include denaturing gradient gel electrophoresis (DGGE), heteroduplex analysis, SSCP analysis and sequencing to identify unknown variation at

a particular locus, and RFLP and allele specific oligonucleotide analysis (ASO) to detect known variation. The SSCP analysis, sequencing, RFLP detection and ASO analysis techniques were used in this project and will be described in detail.

2.2.7.1 Single strand conformation polymorphism (SSCP) analysis

Orita and colleagues (1989a,b) first described PCR-SSCP analysis, a gel-based technique used to detect DNA variation. SSCP analysis was used in this study to screen the exons of *TYR* and *TYRP1* for variation in the DNA sequence in 15 random Negroid and 15 random Caucasoid individuals.

SSCP analysis relies on conformational intrastrand differences in DNA with different sequences. It exploits the fact that PCR generated DNA strands form different secondary conformational structures if the strands have different sequences. The sequence of interest is first amplified using ^{32}P-dCTP to facilitate the detection of it. The PCR product is then denatured into single stranded molecules by heating and electrophoresed vertically through a polyacrylamide gel in non-denaturing conditions. The single stranded DNA molecules take up secondary conformational structures as they run through the gel depending on their base sequence. The differing structures cause the molecules to run through the gel at different rates. The gel is dried and exposed to X-ray film for the detection of band shifts. The variation in the sequences is detected when the samples with band shifts are sequenced.

SSCP analysis works optimally with DNA fragments 300bp or less in size. A study by Shefield *et al* (1993) on the mouse globin gene mutations revealed that the sensitivity of this technique varied considerably with the size of the DNA fragment. The optimal size for SSCP analysis was found to be 150bp and the sensitivity of SSCP decreases with increase in size of the fragment [Shefield *et al*, 1993]. Other studies indicated that in fragments less than 200bp in length, 90% of the variation could be detected, while 80% of variants could be detected in fragments of 300 to 400bp in length [reviewed by Hayashi and Yandell, 1993]. This level of sensitivity was the reason that this technique was chosen for this project. If the PCR product is much larger than 300bp, restriction enzymes are used to obtain suitably sized fragments. Table 2.3 shows the PCR products that were digested with restriction enzymes and the fragment sizes that were obtained.

Table 2.3: Restriction enzymes and SSCP product sizes

Gene	Exon	Restriction enzymes	PCR product size (bp)	Digested fragment sizes (bp)
TYR	1A	*Ava II*	498	309, 64, 124
		Hae III		120, 87, 58, 42, 209
	1B	*Hinf I*	548	233, 315
		Rsa I		349, 199
	2	*Hinf I*	420	264, 156
		Ava II		165, 255
	5	*Bgl I*	522	264, 258
		Hinf I		162, 41, 319
TYRP1	2	*Dde I*	490	300, 85, 105
		Hinf I		281, 209
	3	*Dde I*	416	240, 176
		Alu I		278, 35, 103

Acrylamide is a monomer that must be polymerised for the formation of a gel matrix. This gel is used instead of agarose as the former has greater resolving power. The free radicals that are necessary for polymerisation are provided by ammonium persulfate (APS), which is stabilised by TEMED. A chain reaction is initiated in the presence of the free radicals where the acrylamide monomers are polymerised into chains. In the presence of bisacrylamide the chains become cross-linked to form a gel. The size of the pores of the gel depends on the length of the chains, which is determined by the concentration of acrylamide and the degree of cross-linking [Sambrook J. *et al*, 1989]. The DNA fragments in the polyacrylamide gel separate out according to size and shape, depending on the sequence of bases. Glycerol is an added variable in the gel, which affects the folding of the DNA [reviewed by Hayashi and Yandell, 1993]. It reduces the mobility of the DNA through the gel. Studies have shown that electrophoresis at room temperature (20-25°C) with 5-10% glycerol or at 4°C without glycerol are suitable for SSCP analysis [Hayashi, 1991; Hayashi and Yandell, 1993]. The conformational changes of the DNA are viewed by the autoradiography of the gel.

Gel preparation

Glass plates (41-44cm long, 18.5-20cm wide) were first washed, dried and wiped clean with 100% ethanol. One of them was treated with Gel Slick ™ (FMC), which ensures that the gel does not stick to this plate after electrophoresis. This substance is commercially produced and is

used in place of silicon, which was originally used and replaced because of its toxic nature. The apparatus was set up with 0.3mm spacers to separate the plates and was clamped firmly to ensure that the plates stayed in place. The gel used for SSCP analysis was the mutation detection enhancement (MDE) gel (FMC), which has a polyacrylamide type matrix with high sensitivity to DNA conformational differences [MDE protocol]. Gels both without and with 10% glycerol were used. APS (160µl of 10% stock) and 30µl TEMED were added to the MDE gel solution (Appendix 2), which was immediately poured between the plates and allowed to set for at least one hour. The gel was then set up in a vertical electrophoresis apparatus.

Electrophoresis

An aliquot of the ^{32}P-dCTP labelled PCR product (3µl of undigested PCR product or 2.5µl of each digested product, when two restriction enzymes were used) was mixed with formamide dye (Appendix 2) (7µl or 5µl respectively) and denatured at 100°C for 2.5 minutes. The sample (3µl) was then loaded into the wells of the gel. The gels containing glycerol were electrophoresed at 8W (~800V) overnight (approximately 16 hours), while the gels without glycerol were electrophoresed for 5 hours during the day at 15W (~1200V). The electrophoresis was carried out with a fan blowing on the gels to ensure that they did not over-heat. The gels were then transferred to Whatmann paper and dried for 1 hour in a gel dryer (Biorad M583), at 80°C, connected to a vacuum pump (Savant RT100) and exposed to X-ray film (Cronex or Kodak). The X-ray film was developed to detect the band shifts.

2.2.7.2 Restriction fragment length polymorphism (RFLP) detection

Restriction endonucleases were necessary for the detection of the ROCA S166X mutation, which was found to account for 45% of ROCA chromosomes in a study by Manga *et al* (1997). The S166X mutation occurs in exon 3 of the *TYRP1* gene. It is a C→G base substitution that abolishes a *Mbo1* restriction site. The amplified product of the normal gene is cut to produce 261bp and 155bp size fragments. In homozygous mutant individuals only the 416bp product is observed and in heterozygous individuals all three products are observed (Fig.2.1). The frequency of this mutation was investigated in 48 unrelated random Negroid individuals, 47 random Caucasoid individuals, 17 random San individuals and five Caucasoid individuals with red hair.

Exon 3 of the *TYRP1* gene was first amplified according to the conditions in Table 2.1. The PCR product, 20µl, was incubated with 1µl *MboI* (5U) enzyme and 3µl buffer K (Amersham) for each sample. The samples were made up to 30µl with double-distilled water and incubated at 37°C for 2 hours. Controls for these three genotypes were also included in the screen, to ensure that the results were correct. The products were separated by AGE and visualised under UV light.

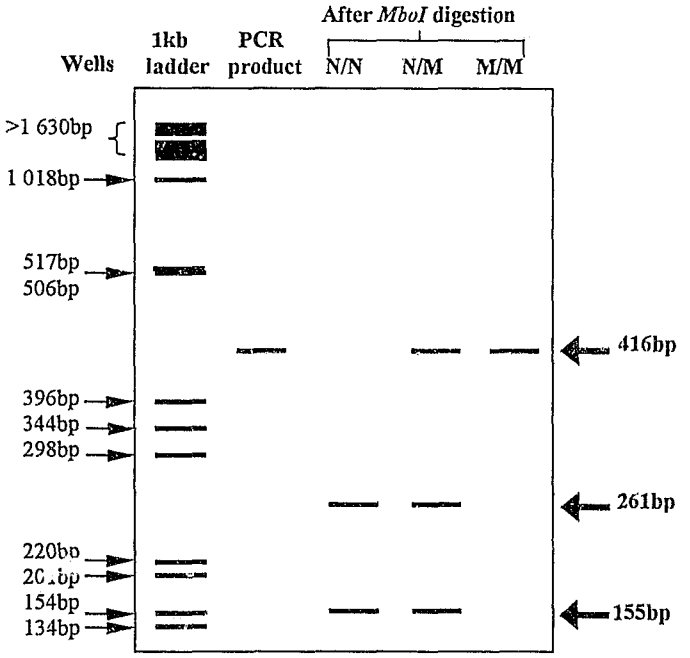


Fig.2.1: S166X detection. A schematic representation of the expected results from *MboI* digestion of exon 3 of the *TYRP1* gene after AGE. The arrows on the left show the different sizes of the 1kb ladder that are used to find the approximate size of the fragment. The arrows on the right of the gel show the sizes of the expected fragments after digestion for a homozygous normal individual (N/N), a heterozygous individual (N/M) and a homozygous mutant individual (M/M).

Exon 3 of the *TYRP1* gene was first amplified according to the conditions in Table 2.1. The PCR product, 20µl, was incubated with 1µl *MboI* (5U) enzyme and 3µl buffer K (Amersham) for each sample. The samples were made up to 30µl with double-distilled water and incubated at 37°C for 2 hours. Controls for these three genotypes were also included in the screen, to ensure that the results were correct. The products were separated by AGE and visualised under UV light.

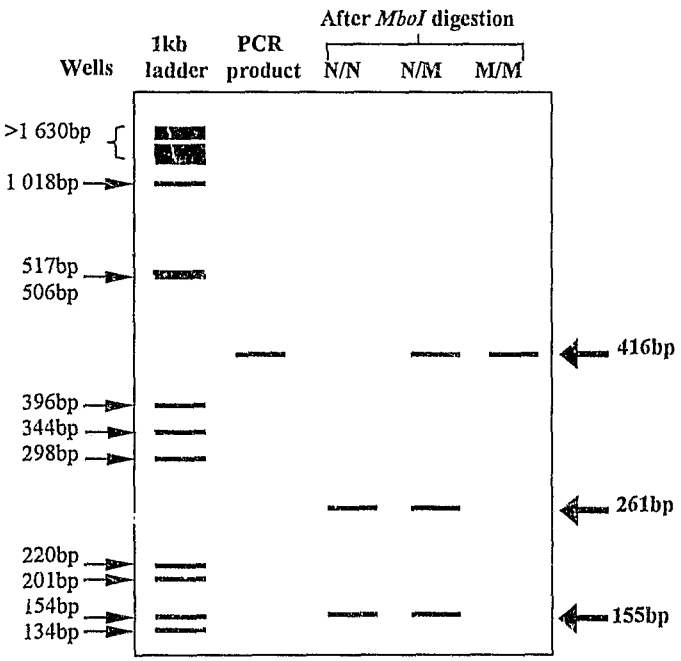


Fig.2.1: S166X detection. A schematic representation of the expected results from *MboI* digestion of exon 3 of the *TYRP1* gene after AGE. The arrows on the left show the different sizes of the 1kb ladder that are used to find the approximate size of the fragment. The arrows on the right of the gel show the sizes of the expected fragments after digestion for a homozygous normal individual (N/N), a heterozygous individual (N/M) and a homozygous mutant individual (M/M).

2.2.7.3 Allele specific oligonucleotide (ASO) analysis

ASO analysis was described by Conner *et al* (1983), where they detected the normal and mutant alleles of the globin gene in sickle cell anaemia (β^A and β^S respectively), with ^{32}P-labelled oligonucleotide probes. This procedure has also been described by Saiki *et al* (1989), using a nonradioactive procedure. In this project ASO analysis was used to detect the 368delA ROCA mutation, where the deletion of an adenine in codon 368 of exon 6 of the *TYRP1* gene is present in the mutant. This mutation was found in 50% of ROCA chromosomes in the study by Manga *et al* (1997). The 368delA was investigated by ASO analysis in 24 random Negroid, 15 random San and 5 red-haired individuals.

Allele specific oligonucleotides (ASOs) are made so that they detect specific sequences in a particular area in a gene. The oligonucleotides are designed complimentary to the part of the sequence where the mutation is known to occur for a mutant and normal allele. These oligonucleotides are then hybridised to DNA samples, which have been linked to a membrane, to detect the presence or absence of these two alleles. In the presence of the mutation the mutant probe (ASO) binds to the DNA on the blot and in its absence only the normal probe binds. However, in a heterozygous individual both probes bind. Two 13mer oligonucleotides (Table 2.4) were designed, specific for the normal and mutant alleles. They were 5'-end labelled with digoxygenin (DIG) (Boehringer Mannheim), a non-radioactive compound, to facilitate detection.

Exon 6 of the *TYRP1* gene was amplified by PCR and bound to the nylon membrane of a dot blot. The ASOs were then hybridised to the blot. The DIG label, now on the DNA, can be detected by either a colour or chemiluminescent method. In this study the chemiluminescent method was used. The hybridisation and detection procedures carried out were as described in the DIG user's guide.

Hybridisation

The blot was first prehybridised with DIG Easy Hyb (DEH) solution, which acts to block areas on the blot that do not contain DNA to prevent the probe from binding to the positively charged membrane. The blots were prehybridised at an appropriate temperature for 2 hours with 20ml/100cm^2 DEH solution in a sealed hybridisation bag. The prehybridisation solution was

discarded and 10ml/100cm^2 blot of hybridisation solution containing either probe (1pmol probe/ml DEH solution) was added. The probe was allowed to hybridise for 3 hours at the optimal hybridisation temperature. The hybridisation solution was then collected and stored at -20°C for later use. The hybridisation temperatures were first chosen according to the melting temperatures (T_m) of the oligonucleotides, which were adjusted to the optimal temperature for each system. Table 2.4 shows the conditions used for the two oligonucleotides. Note that each blot was hybridised by either the normal or the mutant probe, the same blot was not used for both.

Posthybridisation washes

The blot was placed in a wash tray for posthybridisation washes. The wash solutions contained SSC and SDS. The blot was washed twice for 5 minutes at room-temperature in 50ml 2x wash solution and twice for 15 minutes at the appropriate temperature in 50ml 0.5x wash solution with the normal probe (Appendix 2) or 0.1x wash solution (Appendix 2) with the mutant probe (Table 2.4). The blot hybridised with normal probe was washed at the hybridisation temperature (26°C). The blot with the mutant probe needed more stringent wash conditions, at 32°C and a lower concentration of salt.

DIG detection

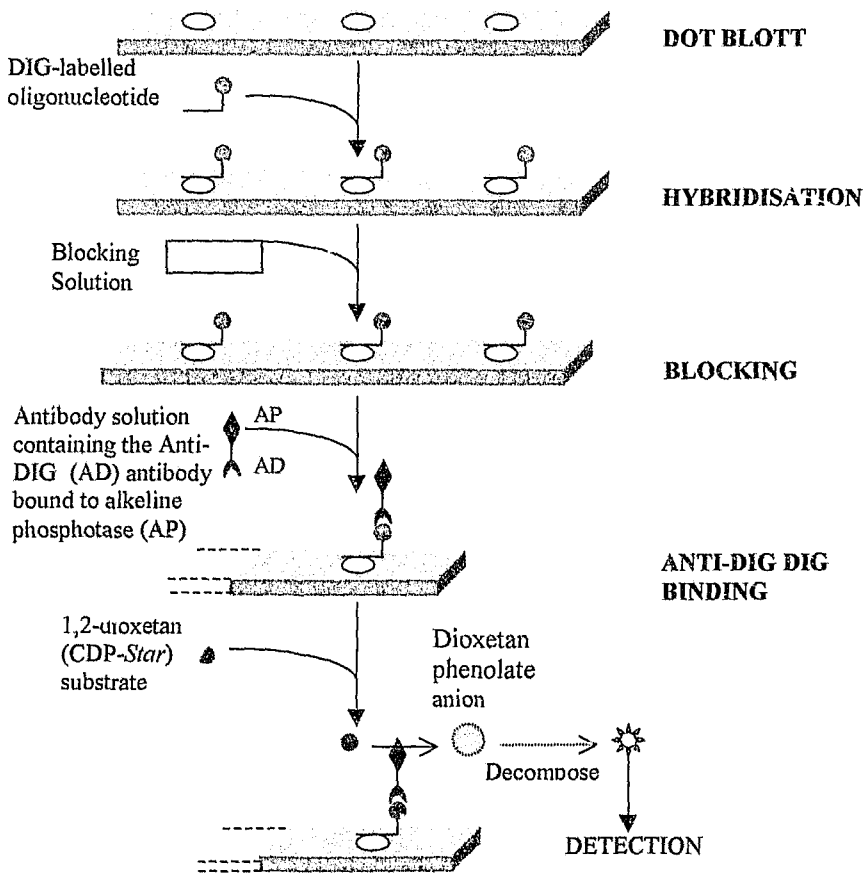
Hybridisation of the probe to the DNA was detected as shown in the diagram in Fig.2.2a. The blot was first transferred to a clean tray. The membrane was rinsed for 1 minute in 40ml 1x wash buffer (10x stock provided), followed by blocking by gentle agitation with 1xblocking solution (10x stock in 1x maleic acid (both provided)) to prevent the antibody from binding non-specifically to the membrane. It was then incubated in 15ml dilute antidigoxygenin antibody solution (1:20 000 of antibody in 1x blocking solution) for 25 minutes. This antibody is linked to an alkaline phosphatase residue, which binds to the DIG tag on the oligonucleotide. The membrane was placed in a new tray and washed twice for 15 minutes in 40ml 1x wash buffer and twice for 5 to10 minutes in 35ml to 40ml more wash buffer to remove excess antibody. These washes were carried out with vigorous agitation. It was then placed in 1x detection buffer (10x detection buffer provided) for 2 minutes to equilibrate. The blot was carefully placed on 500µl CDP-*Star*™ dilution (1:100 dilution of CDP-*Star*™ in 1x detection buffer) and left for

5 minutes for the CDP-*Star*™ to react with the alkaline phosphatase. CDP-*Star*™ is a very sensitive chloro-substituted 1,2-dioxetane chemiluminescent substrate for alkaline phosphotase. The substrate is dephosphorelated to produce dioxetan phenolate anion, which decomposes and emits light at 466nm in a buffer solution (Fig.2.2b) [CDP-*Star*™ protocol]. To detect this light emission the blot was sealed in a fresh plastic bag and exposed for up to 45 minutes to X-ray film (Cronex) and developed.

Table 2.4: ASO hybridisation detection conditions

	Oligonucleotides	
	Normal	Mutant
ASO sequence	5'-ACGGGAAAGTATG-3'	5'-CGGGAAGTATGAC-3'
Prehybridisation temperatures	26°C	30°C
Hybridisation temperatures	26°C	30°C
Posthybridisation wash conditions	Two 5 minute washes in 2X wash solution at room temperature Two 15 minute washes in 0.5X wash solution at 26°C	Two 5 minute washes in 2Xwash solution at room temperature Two 15 minute washes in 0.1X wash solution at 32°C

(a)



(b)

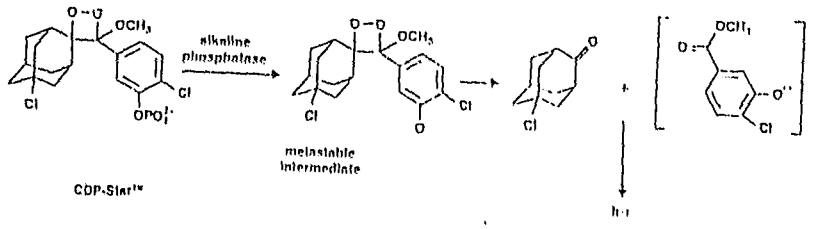


Fig.2.2: ASO detection of the 368delA mutation. (a) An illustration of the process of hybridisation and non-radioactive detection of digoxygenin labelled oligonucleotides by CDP-*Star* and (b) the chemical reaction that results in the fluorescence. [Chemical reaction from CDP-*Star*™ protocol].

2.2.8 Automated DNA sequencing

Automated sequencing is a procedure that can be used to detect variants that are unknown. It was used as a mutation detection procedure for the *MC1R* sequence study, especially as this gene is only 951bp in length. The latter study included 44 Negroid, 17 San and 8 red-haired individuals. This procedure was also used to detect sequence variation in samples that showed SSCP band shifts in *TYR* and *TYRP1* genes.

Maxam and Gilbert, and Sanger first described sequencing in 1977. The latter used an enzymatic method and the former a chemical degradation method; both of which involved the use of radioactivity and four reactions were necessary for each sample, one for each nucleotide. In 1986, Smith and co-workers described the process of automated sequencing using fluorescently labelled primers. Later, Prober *et al* (1987) described the use of fluorescently labelled chain-terminating dideoxynucleotides (ddNTPs) for sequencing. These techniques have the advantage of using fluorescence instead of radioactivity, and are therefore safer. Additionally, only one reaction is necessary for a sample as each of the four ddNTPs is labelled with a different fluorophore, which makes the procedure less labour intensive.

A purified PCR product is used as a template to synthesise fluorescently labelled strands that can be detected by the DNA sequencer. This process of synthesis is known as cycle sequencing, where the reaction goes through many cycles, as in PCR, to produce new strands. During these cycles the primer binds to denatured DNA and the Taq-polymerase enzyme synthesises a new strand using normal dNTPs (deoxynucleotides) and fluorescently labelled ddNTPs. The ddNTPs are chain-terminators that cause the chain formation to stop. The chemical structure of the ddNTP does not allow the enzyme to add any more nucleotides to the chain. Hence, differently sized fluorescent products occur, which can be separated by electrophoresis. The laser beam excites the fluorescence in the chain-terminator-labelled DNA, which is detected by the sequencer via a CCD camera. This information is transferred to the computer, where it is analysed by the appropriate software.

In 1997, Rosenblum *et al* introduced two new types of ddNTPs; dRhodamine and BigDye™ chain-terminators, which have structures that allow for cleaner and more accurate sequencing

results. Fluorescent ddNTPs have the basic structure of the ddNTP linked to a dye via a linker. Both the linker and the dye affect the pattern of termination. The dRhodamine terminators have a propargyl ethoxyamino (EO) linker in-place of the original propargylamino (PA) linker, except ddATP that retains the original PA linker. The EO linker equalises the incorporation rate of the ddNTPs, which results in an even pattern peak. These dyes differ from the conventional rhodamine dyes (R110, R6G, TAMRA and ROX) in structure in that they are dichlororhodamine dyes- named dR110 (on ddG), dR6G (on ddA), dTAMRA (on ddC) and dROX (on ddT); which produce black, green, blue and red signals respectively. The dRhodamine dyes have a narrower emission spectrum, which reduces spectral overlap and, therefore, reduces background noise [dRhodamine terminator cycle sequencing kit protocol]. The BigDye™ ddNTPs are energy-transfer dyes, which have increased signal compared to the other two sets [Rosenblum *et al*, 1997]. They too have the EO linker.

Cycle sequencing

PCR products purified by either of the gel excision procedures or the enzyme cleanup procedure were sequenced. Cycle sequencing is the process that gives rise to amplified, fluorescently labelled sequencing product. One µl primer (3.3pmol/µl) was added to the appropriate amount of clean DNA template (100-200pmol). Four µl or 8µl termination mix and was added to the DNA (for products smaller than 350bp and for products greater than 350bp respectively). The termination ready reaction mixes used were either dRhodamine or BigDye™ (Perkin Elmer). These terminator mixes consisted of the four ddNTPs; the dNTPs; AmpliTaq DNA polymerase, FS; $MgCl_2$ and Tris-HCl buffer, pH 9.0. The sequencing mix was made up to 20µl with double-distilled water. The samples were cycle sequenced in the Hybaid Touchdown or Cetus thermocyclers, where the samples went through the following steps for 25 cycles- 95°C denat. ring for 30 seconds, 50°C for 15 seconds, and 60°C for 4 minutes.

The optimal product length for sequencing is approximately 400bp. As the *MC1R* PCR product was very large (1238bp), six sequencing primers were necessary to obtain the sequence of the whole exon. Fig.2.3 is a schematic representation of the position of the primers in the gene, which are listed in Table 2.5. The 2F and 4R primers produced sequenced products approximately 500bp in size and 6µl of reaction mix was found to be sufficient for the reaction.

Table 2.5: Sequencing primers for the *MC1R* gene

Primer sequence (5' to 3')	Primer direction	Base position	Reference
2F CCCCTGGCAGCACCATGAACT	F	IVS -99 to -79	Box *et al*, 1997
S1 GACAATGTCATTGACGTG	F	349 to 366	Rana *et al*, 1999
S4 TCACCCTCACCATCCTGC	F	722 to 739	Rana *et al*, 1999
2R TGCCCAGGGTCACACAGGACC	R	IVS +53 to +73	Box *et al*, 1997
3R GCGCTGCCTCTTGTGGAG	R	670 to 687	Box *et al*, 1997
4R ATGGAGCTGCAGGTGATC	R	366 to 383	Box *et al*, 1997

The positions of these primers are represented on the schematic diagram in Fig.2.3.

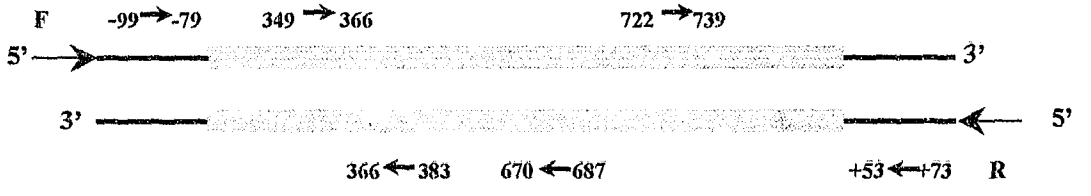


Fig.2.3: *MC1R* primer positions for PCR and sequencing. The PCR primers are indicated in blue, and the forward and reverse primers are shown in red and green respectively. The numbers on either side of the primers are the base positions of the sequencing primers. The coding region is the grey block and the 5' and 3' untranslated regions are the black lines.

Sequencing was also carried out on the *TYR* and *TYRP1* PCR products that showed band shifts in SSCP gels. Variations in the *TYR* SSCPs were observed in exons 1B (548bp), 3 (336bp) and 4 (338bp). 8µl terminating mix was used for the very large exon 1B and 4µl for the rest. Variants in the *TYRP1* SSCPs were found in all seven exons investigated. Exons 2 (490bp) and 3 (415bp) were the only large products and required 6µl termination mix, while the rest of the exons used 4µl mix.

Purification of sequenced product by Centricep columns

Cycle-sequenced product contains unused dNTPs, ddNTPs and primer. The ddNTPs must be removed from the sample, as they may result in a high fluorscent signal, which would interfere with the sensitive detection of the sequencer. Four procedures have been described: Centricep columns, sephadex columns, sodium acetate and magnesium chloride. The four procedures were tested; and the Centricep columns were used as they were found to be the optimal.

The Centricep columns were hydrated with 800µl double-distilled water for at least 2 hours. Excess water was removed by allowing the columns to drip for 15 minutes. They were then centrifuged for 2 minutes at 3 000rpm so as to tightly compact the columns. The samples were then added to the columns and centrifuged again at 3 000rpm for 2 minutes to elute the sequenced product. The samples were then vacuum dried in a spin vacuum apparatus (Savant RT100 vacuum) and the speed vac concentrator (Savant). The dried product was stored at -20°C for no more than 2 weeks, before being analysed on a gel.

Electrophoresis

The 4.3% polyacrylamide gel (40ml) (Appendix 2) was polymerised by the addition of 200µl 10% APS followed by 24µl TEMED. The gel was immediately poured between two glass plates (42cm in length and 35cm wide) separated by 0.2mm spacers set up in the gel pouring apparatus (Perkin Elmer). The gel was allowed to set for at least 1.5 hours. The samples were resuspended in 3µl or 1.5µl dextran-formamide dye (Appendix 2), denatured for 2 minutes at 100°C and loaded on the 32cm gel with a 24 or 36-well comb respectively.

The samples were then electrophoresed with 1XTBE running buffer in an ABI Prism™ 377 automated sequencer at a voltage of 1680V, 50mA current and 150W power. The run module used was 36E-1200, as was the pre-run module. The data were collected by ABI Prism™ 377 collection software version 2.1. The results were analysed by the sequence analysis program, the ABI 377 Prism™ DNA sequencing analysis software, version 2.1.1, MapApp™ ©1985-1995. Finally, an electropherogram was obtained of the DNA sequence of interest. The sequence on the electropherogram was compared to the consensus sequences, which are presented in Appendix 4.

2.3 Statistical analysis

The significance of the variants found in the *MC1R*, *TYR* and *TYRP1* genes was analysed using the following statistical tests.

2.3.1 Allele frequency and haplotype frequency calculations

Allele or variant frequencies were calculated for the three loci investigated. The *MC1R* gene revealed many variants, particularly in individuals with red hair and pale skin. The frequencies calculated in the different population groups were compared to each other. Comparisons were also made for known variants with data from other studies. The allele frequencies in the *TYR* and *TYRP1* genes were, likewise, compared between the Negroid and Caucasoid groups that were studied, and to frequencies reported by other groups.

Allele frequencies (f) were calculated by:

$f = n / nt$ .....equation 2.1

where n = number of chromosomes with particular variant

nt = total number of chromosomes.

Haplotype frequencies were also calculated, as some individuals had more than one variant at a locus, to compare populations at a particular locus. The data were analysed by the Arlequin program (version 1.1). This program has the added advantage of the ability to estimate haplotypic data when the genotypic data are presented with an unknown phase of particular variants. More than one variant was found in several individuals at the *MC1R* and *TYR* loci

and, therefore, haplotype frequencies were calculated for these two loci. The frequencies of the haplotypes were calculated using the Arlequin program (version 1.1), which uses an input file in a particular format, depending on the availability of the data. Examples of the input files for the three loci are presented in Appendix 5.

Several of the individuals at the *TYR* locus were doubly heterozygous, and therefore, the phase was unknown. The most likely haplotypes were deduced using the Arlequin program, with the data in the input file being presented as genotypic data with unknown phase. Arlequin uses a maximum likelyhood method to perform this estimation, assuming Hardy-Weinberg equilibrium.

The *MC1R* locus data were entered as haplotypic data since the phase in each individual could be ascertained using the parents' genotypic data (as explained in section 3.1.1). However, for the San and Caucasoid red-haired individuals analysed for the *MC1R* study, parental data were not available for the determination of phase and the haplotype frequencies were estimated using the program.

2.3.2 The Chi-squared test

The chi-squared test (χ^2) can be used to determine if an observed experimental result fits the theoretical expectations. The χ^2 value obtained is used to estimate how frequently the observed deviation can be expected to occur as a chance event. The χ^2-test can also be used to determine that the alleles in a population are in Hardy-Weinberg equilibrium and to see if a significant difference existed between the populations studied. In this study the χ^2-test was used to compare the Negroid and San population groups with respect to the variants observed in the *MC1R* gene, and to compare the Negroid and Caucasoid populations at the *TYR* and *TYRP1* loci.

The basic χ^2 test is:

$\chi^2 = \Sigma [(o - e)^2 / e]$ .....equation 2.2

where o = the observed value

e = the expected value

(o-e) = the standard deviation in each case

In this study two populations were compared at a particular locus. The χ^2 value obtained indicates a probability (p) value from the χ^2 table, which determines the validity of a null hypothesis, to show that the null hypothesis can be rejected or accepted at either the 5% ($p<0.05$) or 1% level ($p<0.01$). The null hypothesis in the population studies carried out is the hypothesis of no difference between the two populations being analysed, i.e. in equation 2.2 one population is treated as 'expected' and the other as 'observed'. This study made use of the χ^2 program to analyse the data.

2.3.3 The exact test of population differentiation

The exact test of population differentiation examines the hypothesis of a random distribution of individuals between pairs of populations. This project concerns the contribution of *MC1R*, *TYR* and *TYRP1* genes to the differences observed in pigmentation in different population groups. The exact test of population differentiation gives the probability of the extent of the difference between the populations. This test was also carried out on Arlequin, and is analogous to the Fisher's exact test on a two-by-two contingency table. It tests the hypothesis of the random distribution of individuals between pairs of populations. The calculation, therefore, helps to ascertain the part that the locus in question plays in the different pigmentation types observed in the population groups studied. If the probability value is less than the significant value (0.05) the two populations are considered to be significantly different.

2.3.4 Nucleotide and gene diversity

Nucleotide diversity (π) is the number of nucleotide differences per site between two randomly chosen sequences. It can be calculated using the equation [Li, 1997]:

$V(\pi) = V(\Pi)/L^2$

where Π = the average number of nucleotide differences between two sequences randomly chosen from the population

V = variance

L = length of the sequence

The nucleotide diversity was calculated for *MC1R*, *TYR* and *TYRP1* using the Arlequin program, with the Tajima and Nei correction. The correction is necessary to account for the fact that the number of nucleotide changes detected at a site is smaller than the actual number

of substitutions in a site where multiple substitutions have occurred. Tajima and Nei (1984) provide a model where the frequency of substitution of the four nucleotides does not have to be assumed to be equal [Li, 1997]. Complete nucleotide sequences were provided to the Arlequin program (not presented in Appendix 5). This calculation not only gives a statistical value for the diversity of a gene at the nucleotide level, in a certain population, it also provides an estimate of the relative age of the populations [Li, 1997]. Larger nucleotide diversity is usually a good indicator of the age of the population, since more elapsed time allows for the accumulation of more mutations, although other factors may play a role in determining nucleotide diversity, such as selection, population admixture, population bottle-necks, drift and founder effect.

Gene diversity is used as a measure of genetic variability within a population and is defined as the probability that two sequences chosen at random from a population are different, or the expected heterozygosity for a randomly mating populatic [Li, 1997]. It is equivalent to nucleotide diversity at the gene level. This value is not suitable for DNA data as the diversity at the DNA level is very extensive, therefore nucleotide diversity calculations are more appropriate [Li, 1997]. Hence, gene diversity was calculated at the protein level, and only the variants that produced changes at the protein level, i.e. non-synonymous variants, were taken into account and presented in the data input file (not presented in Appendix 5). Gene diversity calculations were also carried out on Arlequin for the *MC1R*, *TYR* and *TYRP1* loci. This program calculates the probability that two haplotypes chosen at random are different in a sample.

2.3.5 The Hardy-Weinberg equilibrium test

The Hardy-Weinberg law states that gene and genotype frequencies will remain in equilibrium in a large population in the absence of mutation, migration and selection. In such a case, when considering two alleles (A and B) in a population, the sum of the frequency of the possible genotypes is 1:

$p^2 + 2pq + q^2 = 1$

where p = frequency of allele A

q = frequency of allele B

The Hardy-Weinberg equilibrium test in the Arlequin program tests the hypothesis that the observed genotypes are a product of a random union of gametes in the population. This program uses a test analogous to the Fisher's exact test on a two-by-two contingency table. The probability value that is obtained indicates whether the null hypothesis should be accepted or rejected. The calculation results in a probability val ... which provides the basis to reject or accept the hypothesis at the 5% level.

2.4 Sequence alignment

The DNA and protein sequences were aligned to sequences from other species to find the conserved sequence positions and to determine the implications of the sequence changes observed in this study. This was carried out with the *MC1R* gene using the DNASTAR program, 1997, MEGALIGN feature. The DNA and amino acid sequences were aligned using the Clustal method. The Clustal algorithm groups the sequences into clusters by examining the distances between pairs of sequences. These are then arranged into pairs and then into groups. This procedure was not carried out with the *TYR* gene as the variants found were previously reporte nd have been analysed by other groups.

3. RESULTS

The *MC1R*, *TYR* and *TYRP1* genes were investigated in individuals with various types of normal pigmentation. Various experimental procedures were carried out to obtain the DNA sequence variants in these different groups of people. These results will be presented and the possible significance pointed out.

Mutation notation is not always consistent between publications. In this dissertation the one letter code for amino acids will be used consistently for mutations in the coding region, for example S47I and not Ser47Ile. Mutations in the intronic sequence will have the following notation: IVS (intervening sequence) number, followed by a "+" or "-" and the number of bases for mutations downstream from the 3' end and upstream from the 5' end of the exon respectively and the base change that was observed. For example IVS4-4G→T.

3.1 Variation at the *MC1R* locus

In order to understand the significance of the role that *MC1R* plays in pigmentation, variation at this locus was investigated by sequencing since it is a small gene with a single exon of less than 1kb in length. The gene was first amplified by PCR, purified by the Wizard-miniprep or QIAquick procedure, sequenced non-radioactively and analysed with sequencing software. Statistical analyses were used to investigate the significance of the variants observed within and between populations.

To analyse the variants the DNA and protein sequences from different mammalian species were aligned, using the MEGALIGN feature of the DNASTAR program. The Clustal method was used. The human *MC1R* DNA sequence (accession number X65634) was aligned to *MC1R* sequences from other species, the sequences of which were also obtained from GenBank. These were: *M. musculus* (mouse) (accession number X65635), *C. familiaris* (dog) (accession number AF064455), *V. vulpes* (red fox) (accession number X90844), *E. caballus* (horse) (accession number X98012), *O. aries* (sheep) (accession number Y13965), and *O. moschatus* (ox)

(accession number Y13956). Human *MC1R* DNA has only 75.5% homology to mouse DNA, and more than 80% similarity with the other species shown here, the highest being 83.4% with sheep *MC1R*. The protein sequences used had accession numbers CAA46588, CAA46589, AAC33737, CAA62349, CAA66641, CAA74298, and CAA74290 respectively.

3.1.1 Sequencing results of *MC1R* in the random groups of Negroid and San

The sequences were compared to the consensus sequence obtained from Rana *et al* (1999). The consensus sequence was modified from that reported by Mountjoy *et al* (1992) to P162R (CCG→CGG) and T90S (ACG→AGC), changes which were also observed by Chhajlani and Wikberg (1992). Additionally, a Q163R (CAA→CGA) difference was observed to the sequence published by the latter group. Sequencing analysis of *MC1R* in Negroid and San individuals revealed many changes in the coding region, both synonymous and non-synonymous mutations.

A total of 22 mother-child pairs from the Negroid population were sequenced, as well as 17 random individuals from the San population. In the Negroid group the results in the mothers were used to find the phase of the variants observed in the children. This was possible in 15 cases. The mothers' and children's haplotypes were deduced by assuming that no crossover had taken place during gametogenesis. In cases where the inheritance pattern was not ambiguous the mother's chromosome that had not been inherited by the child was used as the third chromosome. Three random chromosomes could then be included in the analysis for many of the families. In two cases it was difficult to deduce the phase of the child's variants from the mother's data. In these cases, the relevant areas of the father's *MC1R* gene were sequenced to deduce phase. Hence, a total of 59 random Negroid chromosomes could be used for analyses and interpretation. 32 San haplotypes were unambiguous as the 16 individuals either presented with heterozygosity at one locus or no variation in the sequence. One individual in the San group was heterozygous at two loci in the *MC1R* gene and the phase of the alleles could not be determined.

Two non-synonymous or missense mutations (S47I and F196L) and seven synonymous or silent mutations were observed in the Negroid group. The San had relatively fewer variants; four in total (one non-synonymous (L99I) and three synonymous). The mutations are listed in Table 3.1.

All these DNA sequence variants, except T314T, have not been described previously. The non-synonymous mutations are of particular interest as they may be functionally significant. Appendix 4 contains the consensus *MC1R* sequence that was used in this study.

Sequencing using the automated sequencer is much less labour intensive than the manual method. The sequencing program analyses the results and produces an electropherogram, which is edited by scrutinising for possible ambiguities. Fig.3.1 shows an example of a set of sequencing results comparing a normal sequence (a), a heterozygous sequence (b) and a homozygous mutant sequence (c) of the T314T variant, the most frequent of the variants observed in the Negroid group. The sequences were read in both the forward and reverse directions to confirm the results.

3.1.1.1 Non-synonymous mutations

The most frequent non-synonymous mutation in Negroid individuals was F196L, which has not been reported previously. Phenylalanine and leucine are both amino acids with non-polar side chains, and this mutation is, therefore, unlikely to have an influence on the protein structure due to charge. However, since it occurs within the fifth transmembrane domain it may affect the binding of the receptor to the melanocyte membrane. The two amino acids are of different molecular weights, with leucine being lighter than phenylalanine. Hence, the change in weights at that position may have an effect on the positioning of the protein in the melanocyte membrane, and hence its function. This change was observed at a frequency of 0.119 in the Negroid group (7/59 chromosomes), but was not observed in the San (Table 3.2). Table 3.3 presents the data according to language groups. Only the mothers of the Negroid group were included, as some of the children were the result of marriages of individuals from different language groups and could not be placed in any particular group. A comparison of the language groups within the Negroid population revealed that the F196L mutation occurred mainly in the South Sotho-speakers (Table 3.3). These individuals, as other Sotho-speakers, are known to practice consanguineous marriages. Although this fact should not affect the frequency of this mutation in this Negroid group, it would maintain this variant in this group of individuals and increase the occurrence of homozygotes.

The other two non-synonymous mutations observed were S47I in the Negroid and L99I in the San. They both occurred in heterozygous form in one individual each. S47I had a frequency of 0.017 in the Negroid group. This variant occurs in the first transmembrane domain (Fig.3.2) and, therefore, this change may affect the binding of the protein to the melanocyte membrane. Serine has an electrically neutral polar side chain and isoleucine has a non-polar side chain. This change, therefore, may have an effect on the structure of the MC1R protein, which may also be affected by the change in molecular weight at position 47. L99I had a frequency of 0.029 in the San. This variant occurs in the second transmembrane domain of the MC1R protein (Fig.3.2). Leucine and isoleucine have non-polar side chains. Hence, the change may not have had an effect on the structure of the protein due to charge and weight (as both these residues are nearly the same molecular weight), but may have had an effect on the function and its binding to the membrane. This cannot be determined until functional studies are carried out.

Fig.3.2 is an illustration of the MC1R protein and its folding in the melanocyte membrane. The positions of the synonymous and non-synonymous mutations in the *MC1R* gene are indicated in different colours for the Negroid (green), San (yellow) and Caucasoid individuals with red hair (red).

The alignment of the human DNA and protein sequences with that of other mammalian species showed the regions that have been conserved at the *MC1R* locus between them. The alignment results are presented in Appendix 5. The base positions in the DNA sequence and the amino acids in the protein that were changed in the Negroid and San are indicated. The DNA and protein sequence alignment results show that the position of the three non-synonymous changes that were observed in the Negroid and San were highly conserved between species, suggesting that the changes may be functionally significant.

3.1.1.2 Synonymous mutations

The Negroid and San groups shared two synonymous mutations, V265V and T314T, out of the total eleven synonymous variants that were observed in them. The frequency of V265V was higher in the San than the Negroid, 0.059 (2/34) and 0.017 (1/59) respectively (Table3.3). The T314T synonymous variant was the most frequent variant in the Negroid individuals (0.424;

25/59) and was described previously in African individuals by Rana *et al* (1999), where it had an allele frequency of 0.42 (21/50). It has also been reported in some Asian and Caucasoid individuals [Box *et al*, 1997; Rana *et al*, 1999]. It was found in the San with a frequency of 0.059 (2/34). The most frequent variant in the San was L50L with a frequency of 0.088 (3/34).

DNA sequence alignments were done to see if the bases were conserved between species in the positions of synonymous mutations. The alignment results show that most of the positions at which base changes were seen were highly conserved between species, with the exception of nucleotides 309, 699 and 942. There was a C→T transition at 309. This base was also a T in the ox and sheep *MC1R* sequences and was an A in the dog and fox and a G in the mouse, suggesting that this base change is unlikely to be of functional significance. At position 699 in the mouse an A was observed in position of the G, as observed in the variant in one Negroid individual. The A→G transition at 942 is also unlikely to be significant as the mouse, ox and sheep each have a G at this position.

3.1.2 *MC1R* DNA variation in red-haired individuals

The red hair and pale skin phenotype has been found by many studies to be linked to mutations in the *MC1R* gene (Valverde *et at*, 1995; Box *et al*, 1998; Rana *et al* 1999). In this project, a total of eight red-haired individuals were studied, two of who were related (mother and daughter). Nine non-synonymous mutations were found in the eight individuals. Four of these mutations have not, to my knowledge, been reported previously. They are S83P, Y152X, A171N and P256S and their positions in the protein, as well as the remaining five mutations that were observed are shown in Table 3.4. The mutations are also represented in Fig.3.2. All individuals were either homozygous for one mutation or compound heterozygotes for two different mutations. None of them had a normal *MC1R* gene.

Eight of the nine mutations are missense mutations, while one (Y152X) is a nonsense mutation that results in a truncated protein. The nonsense mutation was found in heterozygous form with the V60L mutation, possibly on the other copy of the gene. The most frequent mutation in the red-haired individuals was R151C, which had a frequency of 0.286 (4/14) in the random group

excluding the mother of the mother-daughter pair (Table 3.4). The electropherogram illustrating this frequent variant in the Caucasoid individuals with red hair is presented in Fig.3.3.

Serine and proline have polar and non-polar side chains and do not contribute to the charge of the protein. Hence, the S83P and P256S changes may not significantly affect the structure of the MC1R protein because of charge. However, proline is heavier than serine, and the change in size and shape may have an effect on the protein's structure and function. Alanine and asparagine are also not known to contribute a charge as alanine is non-polar and asparagine is polar, and therefore the A171N change probably does not produce a marked change in the structure of the protein. However, asparagine is much heavier than alanine and the change in size and shape may affect the structure and function of the protein.

Both amino acids of the previously reported V60L variant are non-polar and therefore the size and shape changes may have an effect on the structure. This variant has also been noted in individuals without red hair [Box *et al*, 1997; Koppula *et al*, 1997]. Argenine and histidine are basic amino acids and have a positive charge at neutral pH. The R142H mutation may, therefore, have little effect on protein structure. R151C and R160W produce a basic (positive at neutral pH) to an electrically neutral polar side chain change. The D294H variant produces an acidic, which is negative charge at neutral pH, to a basic amino acid change. The change in charge of these three variants is likely to have an effect on the structure of the MC1R protein and its folding and alignment in the melanocyte cell membrane, and may affect the function of this receptor in the signalling and activation of cAMP.

At the DNA level, the base changes that occurred at positions 456, 478 and 512 (at amino acids 152, 160 and 171 respectively) were not conserved in the mouse, and position 512 was not conserved in the horse as well. The protein alignment results were used to investigate the mutations that were observed in these red-haired Caucasoid individuals (Appendix 5). Most positions at which amino acid changes were observed in this study were completely conserved between the species, suggesting that they may be functionally significant. The exception is A171N, where the mouse and horse have a valine at this position. The positions at which these changes occurred have been indicated.

Table 3.1: ***MC1R*** **variants seen in normally pigmented Negroid and San individuals**

Variant	Position in coding sequence	Base change	Position in protein	Number of chromosomes		Reference
				Negroid	San	
S47I*	140	AGC→ATC	First transmembrane domain	1/59	0	Present study
L50L	150	CTG→CTA	First transmembrane domain	0	3/34	Present study
L99I**	295	CTC→ATC	Second transmembrane domain	0	1/34	Present study
A103A	309	GCC→GCT	Second extracellular loop	1/59	0	Present study
L106L	318	CTG→CTA	Second extracellular loop	5/59	0	Present study
I168I	504	ATC→ATT	Fourth transmembrane domain	1/59	0	Present study
F196L*	586	TTC→CTC	Fifth transmembrane domain	7/59	0	Present study
Q233Q	699	CAG→CAA	Third intracellular loop	1/59	0	Present study
V265V	795	GTC→GTG	Sixth transmembrane domain	1/59	2/34	Present study
F300F	900	TTC→TTT	Seventh transmembrane domain	2/59	0	Present study
T314T	942	ACA→ACG	COOH terminal region	25/59	2/34	Box *et al* (1997) Rana *et al* (1999)

The chromosome numbers are shown as the Negroid children were analysed along with the mother's chromosome that was not inherited by the child, where the inheritance pattern was clear.

* The non-synonymous mutations observed in the Negroid group.

** The non-synonymous mutation observed in the San group.

Table 3.2: *MC1R* variant frequencies and Chi-square calculations comparing the Negroid and San populations

Variant	Number of chromosomes with variant		Frequency of variants		Chi-square calculations	
	Negroid (59)	San (34)	Negroid	San	Chi-square	Probability
S47I	1	0	0.017	0	0.583	0.445
L50L	0	3	0	0.088	4.938	0.026*
L99I	0	1	0	0.029	1.753	0.186
A103A	1	0	0.017	0	0.583	0.445
L106L	5	0	0.085	0	3.045	0.081
I168I	1	0	0.017	0	0.583	0.445
F196L	7	0	0.119	0	4.362	0.037*
Q233Q	1	0	0.017	0	0.582	0.445
V265V	1	2	0.017	0.059	1.212	0.271
F300F	2	0	0.034	0	1.109	0.292
T314T	25	2	0.424	0.059	13.94	0.0002**

* There is significant difference at the 5% level.

** There is significant difference at the 1% level.

Table 3.3: Chromosomes with *MC1R* variants in the different language groups of the Negroid mothers

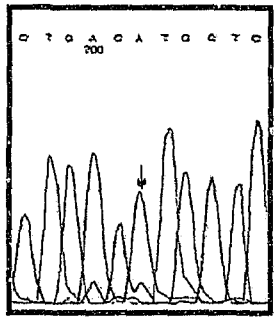
Variant	North Sotho	South Sotho	Tsonga	Tswana	Xhosa	Zulu	Total
S47I*	-	-	-	-	-	-	-
A103A	1/10	-	-	-	-	-	1/44
L106L	1/10	1/16	-	1/4	-	1/6	4/44
I168I	1/10	-	-	-	-	-	1/44
F196L	-	4/16	-	-	-	-	4/44
Q233Q	1/2	-	-	-	-	-	1/44
V265V	-	-	-	1/4	-	-	1/44
F300F	-	2/16	-	-	-	-	2/44
T314T	5/10	5/16	1/2	2/4	1/2	4/6	18/44

*The S47I variant was observed in one child of a South-Sotho mother, but was absent in the mother. The father was listed as Sotho.

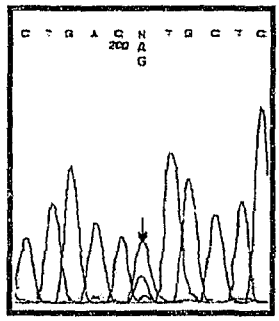
The data are presented as number of chromosomes with the variant out of the total number of chromosomes for a particular language group.

The data from the children were excluded from this table as some were the result of mixed marriages of different Negroid language groups.

(a)



(b)



(c)

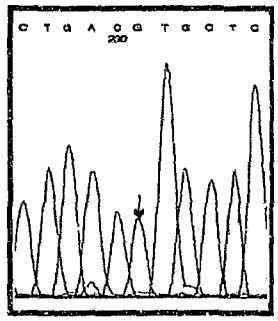


Fig.3.1: Sequencing results. The electropherogram illustrates the most frequent variant in the Negroid individuals, T314T (ACA→ACG). An example of a homozygous normal sample (a), a sample heterozygous for the mutation (b) and homozygous mutant sample (c) are shown. The arrow indicates the position of the variant. The bases for each peak are shown above the peaks. An N is shown at the mutation site in a heterozygote.

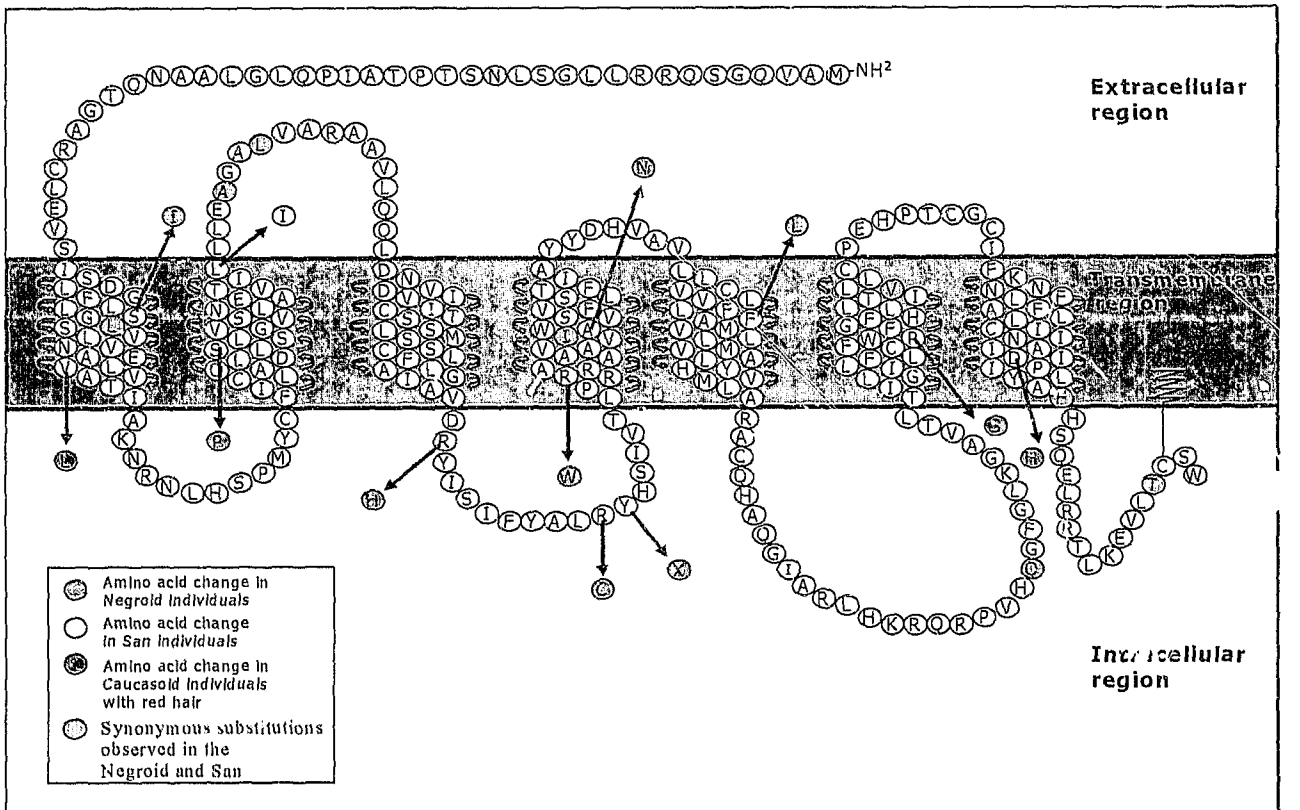


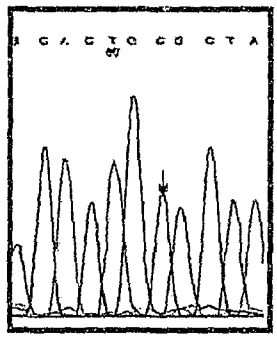
Fig.3.2: Positions of synonymous and non-synonymous mutations observed in the *MC1R* gene. The diagram shows the seven transmembrane domains of the MC1R protein. The non-synonymous mutations observed in the DNA sequences of the Negroid, San and Caucasoids with red hair have been indicated. The synonymous mutations that were observed in the Negroid and San are shaded in pink.

Table 3.4: *MC1R* variants seen in individuals with red hair and pale skin

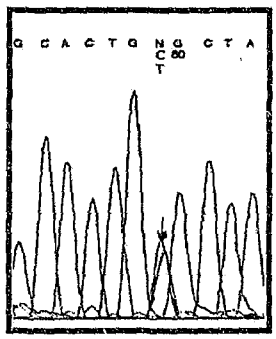
Variant	Position in the coding sequence	Base change	Position in protein	Number of chromosomes *	Variant frequencies	Reference
V60L	178	GTG→TTG	First transmembrane domain	1/14	0.071	Sturm *et al* (1997)
S83P**	247	TCG→CCG	Second transmembrane domain	1/14	0.071	Present study
R142H	425	CGC→CAC	Second intracellular loop	1/14	0.071	Sturm *et al* (1997)
R151C	451	CGC→TGC	Second intracellular loop	4/14	0.286	Sturm *et al* (1997) Rana *et al* (1999)
Y152X**	456	TAC→TAA	Second intracellular loop	1/14	0.071	Present study
R160W	478	CGG→TGG	Fourth transmembrane domain	3/14	0.214	Present study
A171N**	512	GCC→GAC	Fourth transmembrane domain	1/14	0.071	Present study
P256S**	766	CCC→TCC	Sixth transmembrane domain	1/14	0.071	Present study
D294H	880	GAC→CAC	Seventh transmembrane domain	1/14	0.071	Valverde *et al* (1995) Sturm *et al* (1997)

* The daughter of the mother-daughter pair was included in this table to present all the variants in a group of random individuals with red hair.
** The novel mutations.

(a)



(b)



(c)

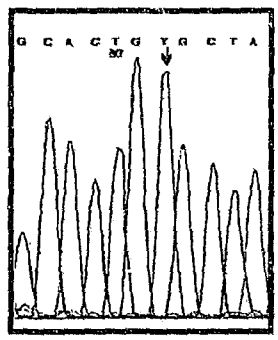
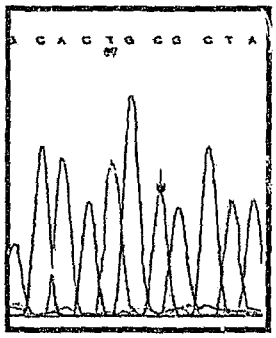
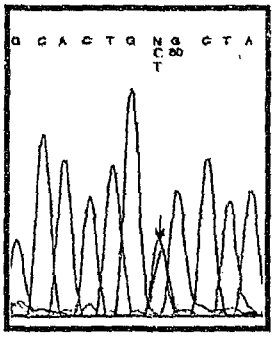


Fig.3.3: Sequencing results. The electropherograms show the most common variant in *MC1R* observed in Caucasoid individuals with red hair, R151C (CGC→TGC): homozygous normal (a), heterozygous for the variant (b) and homozygous for the variant (c).

(a)



(b)



(c)

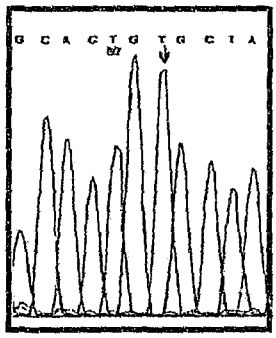


Fig.3.3: Sequencing results. The electropherograms show the most common variant in *MC1R* observed in Caucasoid individuals with red hair, R151C (CGC→TGC): homozygous normal (a), heterozygous for the variant (b) and homozygous for the variant (c).

3.1.3 Statistical analysis of mutation frequencies

The data were analysed by different statistical tests to gauge their significance between the different populations studied. χ^2 analysis was carried out to compare the Negroid and San populations, to see how significantly different these two African populations are at the *MC1R* locus. The computer program Arlequin was used for most of the analyses. It is convenient to use in cases where haplotypic and genotypic data (with alleles having known or unknown phases) need to be analysed. This program provides a wide range of statistical tests. The tests that were used here were the exact test of population differentiation, gene diversity, nucleotide diversity and Hardy-Weinberg equilibrium.

3.1.3.1 Variation between the Negroid and San population groups

MC1R allele frequencies for the Negroid and San are presented in Table 3.2. The Negroid and San population groups were compared to determine whether the *MC1R* gene plays a significant role in the varying pigmentation observed in these two African population groups. A total of 59 chromosomes (random) were used from the Negroids and 34 chromosomes (random) from the San.

The data were examined to identify significant differences of individual alleles between the two groups. χ^2 analysis was used to test the *MC1R* variants between the Negroid and San groups and these results are presented in Table 3.2. The L50L variant was found in the San alone and showed a significant difference between the San and Negroid populations at the 5% level (p=0.026). The F196L variant is very common in the Negroid and was not found in the San. This variant was also significant at the 5% level (p=0.037). The most significant difference between the San and the Negroid groups was with the T314T variant (p=0.0002).

The exact test of population differentiation tests the random distribution of gametes within the populations to give a probability of the difference between two populations at the locus. The program used by Arlequin uses a method analogous to the Fisher's exact test. This test can be carried out with either haplotypic or genotypic data. Haplotypic data were used to compare 59 chromosomes from the Negroid and 32 chromosomes from the San (the phase could not be

determined in one San individual) (Appendix 5). The Negroid and San individuals are significantly different ($p<0.0001$) at the *MC1R* locus.

Haplotype frequencies were calculated by the Arlequin program and, as the data were presented as haplotypes, the procedure was fairly straight froward. The most common haplotype in both groups was that of the consensus sequence, with a frequency of 0.385 in the Negroid and 0.812 in the San. The haplotype containing the T314T variant with no additional changes was the next most common with a frequency of 0.31. All other haplotypes had frequencies less than 0.09. Grouping the chromosomes into haplotypes revealed that the L106L variant was always accompanied by the T314T variant. The consensus haplotype and those with variants at positions 314 or 265 alone were the only three shared haplotypes between the two groups.

Gene diversity tests the heterozygosity at the locus in the populations and provides a probability that two alleles chosen at random from the population are different from each other. This method of analysis was done on protein data, as at the DNA level genetic variation is very extensive. Hence, only the variants in the sequence that produced a change in the protein were investigated and used to calculate gene diversity. This was calculated for the Negroid and San haplotype data. The gene diversity at the *MC1R* locus in the Negroid group studied was high, with a value of 0.242 (+/- 0.068). It was lower in the San, with a value of 0.058 (+/- 0.055).

Nucleotide diversity is the number of nucleotide differences between two randomly chosen sequences. Nucleotide diversity, carried out with sequence data, was found to be 0.12% in the Negroid and 0.05% in the San. Hence, there is more diversity in this gene at the nucleotide level in the Negroid group. The larger diversity values in the Negroid population are compatible with findings at other loci, suggesting that this group is ancient and has not undergone a severe bottleneck and has maintained a large effective population size throughout its history. The San, who are known to be an ancient people were found to have lower diversity at this locus and others at the genetic level [reviewed in Nurse *et al*, 1985; Ramsay *et al*, 1988]. This is likely due to a small effective population size and genetic drift.

Hardy-Weinberg equilibrium is an important requirement to meet when considering population studies, to ensure that a suitable population sample was chosen for analysis. The

Arlequin program tests the hypothesis that the observed diploid genotypes are the product of a random union of gametes. Hardy-Weinberg equilibrium could not be calculated for the Negroid and San with the haplotypic data as only one allele per sample was provided. Hence, the Hardy-Weinberg equilibrium status of the Negroid group of children and of the San was investigated, with the data presented as genotypic data (Appendix 5). The Negroid population sample was in Hardy-Weinberg equilibrium at all eight sites (p=0.463 +/- 0.002). The San group of individuals was also in Hardy-Weinberg equilibrium (p=1).

3.1.3.2 Caucasoid individuals with red hair

The allele frequencies were calculated according to the total number of random chromosomes studied. Fourteen chromosomes were considered when the statistical analysis was carried out. This analysis was carried out with the Arlequin program for genotypic data, with unknown phase (Appendix 5), because it was not possible to determine whether the mutations observed were on two different chromosomes or the same chromosome.

Haplotype frequencies were estimated using the Arlequin program. In this case, where gametic phase was unknown the program uses an algorithm that calculates the maximum-likelihood haplotype frequencies. The haplotype that had a high frequency was the normal consensus sequence. However, it must be noted that it is likely that this haplotype was not present at all, as the variants observed may have been on each chromosome in all the cases presented in this study.

Gene diversity was found to be 0.912 +/- 0.059, which is much greater than the Negroid and San populations at this locus. This shows a clear association between variation at this locus and the red hair and pale skin phenotype.

Nucleotide diversity was 0.19% in the Caucasoids with red hair, almost double that found in the Negroids. This further supports the association between the red hair and pale skin phenotype and variation at the *MC1R* locus.

3.1 Variation at the *TYR* locus

The *TYR* gene exons were amplified for SSCP detection by PCR with ^{32}P-dCTP to label the amplified product. The PCR product was electrophoresed on MDE gels containing 10% glycerol and gels with no glycerol. Variants observed on the SSCP gels were sequenced by the automated procedure. The results were then used for statistical analysis. A total of 15 random Negroid and 15 random Caucasoid individuals were studied to find significant variation between these two groups at the *TYR* locus.

3.2.1 SSCP variants between Caucasoid and Negroid individuals

The band patterns were different between the two different gels for all systems. In most cases gels containing glycerol showed more band shifts than the gels without glycerol. Exons 1B and 4 of the *TYR* gene had SSCP variants with distinctive differences between the Negroid and Caucasoid individuals. Fig.3.4a and b show the SSCP results for these two exons. The arrows on the right side of the gel point out the band shifts that were observed and the numbers at the top indicate the individuals that were studied from the Negroid and Caucasoid population groups.

3.2.2 Sequencing of variants

It is clear from the SSCP gels that there are differences in the banding pattern between the Negroid and Caucasoid individuals at the *TYR* locus. One individual presenting with each type of change in band position was sequenced, as well as a control individual with no band variation. The coding region of the consensus sequence for *TYR* used in this study is presented in Appendix 4. This sequence was obtained from GenBank (accession number M60296) [Giebel *et al*, 1991d]. The only difference to the sequence from GenBank was an A→G change 10 bases 3' of exon 1 (IVS+10A →G) (Table 3.5). Since the G was present in all sequences it was taken to be a part of the consensus sequence in this study. To my knowledge, this variant has not been reported in other studies.

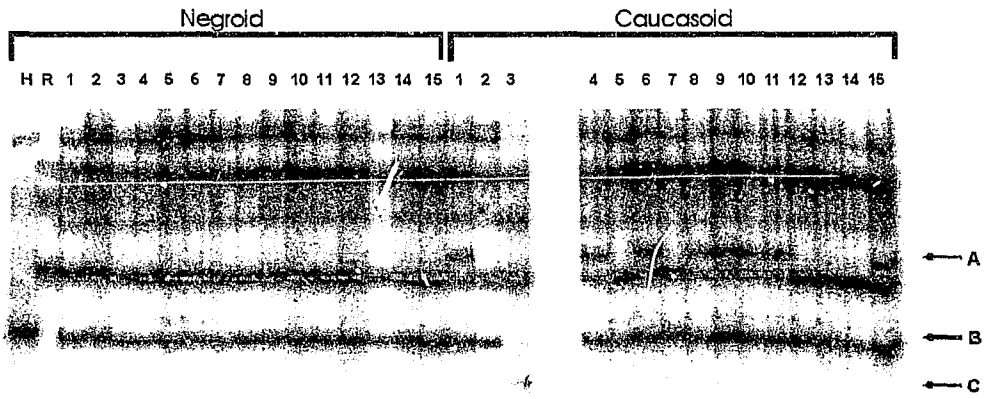


Fig.3.4a): Autoradiograph of the SSCP gel on exon 1B of the *TYR* gene. The results were obtained by electrophoresis of samples through a MDE gel containing 10% glycerol. 15 Negroid samples and 15 Caucasoid samples were investigated. The PCR product of exon 1B is 548bp in length and was digested with restriction enzymes *Hinf I* and *Rsa I*. The product of each digest of one Negroid sample was electrophoresed to observe the fragments produced, labelled H (*Hinf I*) and R (*Rsa I*) respectively. The rest of the samples contain the product of both digests in equal proportion. The arrows on the right hand side indicate the band shifts that were noted. 'A' was present in Caucasoid individuals 1, 4, 6, 8, 9, 10, 11, and 15; 'B' (a shift upwards) was observed only in Caucasoid individual 7 (although not very clear in this copy); and 'C' was present in Caucasoid individual 3. One individual with band 'A', individual 7 with band 'B' (Caucasoid) and individual 3 with band 'C' (Caucasoid) were sequenced, as well as a Negroid individual with no band shifts. Although bands A and B revealed variants (S192Y and R217W respectively), the sample with C had a normal DNA sequence and may have been an artefact or the sample may not have digested completely.

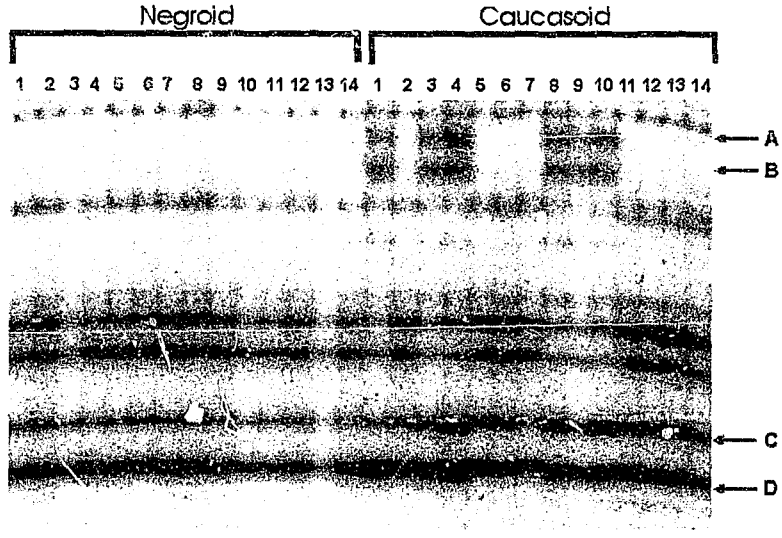


Fig.3.4b): The autoradiograph of the SSCP gel on exon 4 of the *TYR* gene. The samples were electrophoresed in a 10% glycerol MDE gel. Note that 14 Negroid and 14 Caucasoid individuals were investigated here. The arrows indicate the band shifts that were noted. All individuals with bands 'A' and 'B' had extra bands 'C' and 'D' just below the normal band that is seen in the Negroids and the rest of the Caucasoids (not very clear in this reproduction). These band shifts were found in Caucasoid individuals 1, 3, 4, 8, 9, and 10, one of whom was sequenced and found to be heterozygous for the R402Q variant.

The results are summarised in Table 3.5. Two main variants were found: the S192Y mutation in exon 1B and R402Q mutation in exon 4. They had a high frequency in the Caucasoid group studied. These have been described previously as polymorphisms in all populations, except in Orientals [reviewed in Spritz and Hearing, 1995]. Although the S192Y change may have no functional significance, R402Q has been found to result in a temperature sensitive protein [Tripathi *et al,* 1992a].

The S192Y and R402Q polymorphisms were observed to occur in the same individuals in five (Caucasoids 1, 3, 8, 9 and 10) of the nine individuals that had these variants. In these cases the mutations may have occurred on the same chromosome. Three had only the S192Y mutation in heterozygous form and only one individual had the R402Q mutation alone. The frequency of the S192Y polymorphism was reported to be 0.48 and that of R402Q was reported to be 0.15 in all populations except Orientals [reviewed in Spritz and Hearing, 1995]. The S192Y variant was found to have a frequency of 0.267 (8/30) in the Caucasoids in this study and R402Q was found to have a frequency of 0.214 (6/28) in the Caucasoids (Table 3.6). Both were absent in the Negroid group. Argenine is positively charged and glutamine is negatively charged, which may result in a change in the structure and function of tyrosinase. Serine and tyrosine, on the other hand, are both polar, and even though charge may not have played a role the size and shape the change may affect the structure and function of tyrosinase.

One tyrosinase mutation, R217W, known to be associated with OCA1 was observed in exon 1B in Caucasoid individual 7. R217W was described as a rare mutation in Caucasoid OCA1 individuals with a prevalence in their study of <0.02 in Caucasians [Tripathi *et al,* 1992b]. In the present study it has a frequency of 0.033 (1/30) in normally pigmented Caucasoid individuals (Table 3.6). The changes observed in the *TYR* gene and their positions in the protein are illustrated in Fig.3.5.

Table 3.5: *TYR* gene variants found in random normally pigmented Caucasoid and Negroid individuals

Exon	Base change	Position in coding sequence	Variant	Position on protein	Number of chromosomes Caucasoid	Negroid	Reference
1A	None	-	-	-	-		-
1B	TAT→TCT	575	S192Y	Cu-A binding site	8/30	-	Giebel and Spritz (1990)
	CGG→TGG	649	R217W	Cu-A binding site	1/30	-	Tripathi *et al* (1992)
	a→g 10th base after exon		IVS1+10A→G	In 5' donor site	30/30	30/30	Present study
2	None	-	-	-	-		-
3	None	-	-	-	-		-
4	CGA→CAA	1205	R402Q	Cu-B binding site	6/28	-	Tripathi *et al* (1991)
5	None	-	-	-	-		-

Table 3.6: *TYR* gene variant frequencies and Chi-square analysis comparing the random Negroid and Caucasoid groups

Variants	Number of chromosomes Caucasoid	Negroid	Variant frequencies Caucasoid	Negroid	Chi-square calculations comparing Negroid to Caucasoid Chi-square	Probability
S192Y	8/30	0	0.267	0	9.231	0.002
R217W	1/30	0	0.033	0	1.016	0.313
R402Q	6/28	0	0.214	0	6.72	0.01

3.2.1 Analysis of variants and within and between population comparisons

χ^2 analysis was carried out to investigate the difference between the Negroid and Caucasoid populations at the *TYR* locus. The Arlequin program was used to carry out the exact test of population differentiation, to calculate possible haplotype frequencies, gene diversity and nucleotide diversity, as well as to test that the populations were in Hardy-Weinberg equilibrium. The data were presented as genotypes with unknown phase, as the phase of the variants could not be ascertained in cases that had more than one variation in heterozygous form.

The allele frequencies were calculated for the variants and are presented in Table 3.6, which also shows the χ^2 values comparing the Negroid and Caucasoid populations. The two common variants, S192Y and R402Q, are significantly different between the two populations, p=0.002 and p=0.01 respectively. The R217W mutation is not significantly different (p=0.313) between the populations. The exact test for population differentiation, using the Arlequin program, showed that the Caucasoid and Negroid groups were very significantly different at the *TYR* locus (p<0.0001). The input files are presented in Appendix 5. Note that for this and other tests using Arlequin the individuals that were not analysed by SSCP and sequencing for exon 4 (Negroid individual 15 and Caucasoid individual 15) were not included in the data input file.

The haplotype frequencies were calculated by the Arlequin program, which took into account all the possible haplotypes based on the genotypes presented in the input file. In a case such as this, where gametic phase is unknown, the program uses an algorithm that calculates the maximum-likelihood haplotype frequencies. In *TYR*, some individuals were heterozygous for more than one variant and the phase could not be deduced without having access to parent or offspring samples. The most common haplotype in the two populations was the consensus haplotype and this was present in every Negroid individual, while in the Caucasoid population it was the most frequent with a frequency of 0.679. The next most frequent haplotype in the Caucasoids was that which had both common variants observed in this group (S192Y and R402Q), with a haplotype frequency of 0.179. The phase of the two variants is unknown, and it is possible that in some individuals they occur on two separate chromosomes.

The gene diversity was 0.00 for the Negroids, as would be expected for a monomorphic locus, and was 0.582 (+/- 0.101) in the Caucasoids. Nucleotide diversity was found to be 0.05% in the Caucasoids and 0.00% in the Negroids at the *TYR* locus.

Hardy-Weinberg equilibrium was investigated in the Caucasoid and Negroid populations. Exact p=1.00 in the Caucasoid group. Hardy-Weinberg equilibrium could not be calculated for the Negroid group, because the locus was monomorphic. Only one allele was present in this population. It is possible that if a larger group of Negroid individuals were studied, a very rare variant may have been found, and this calculation would have been carried out. Hence, population size may have been the limiting factor in this case.

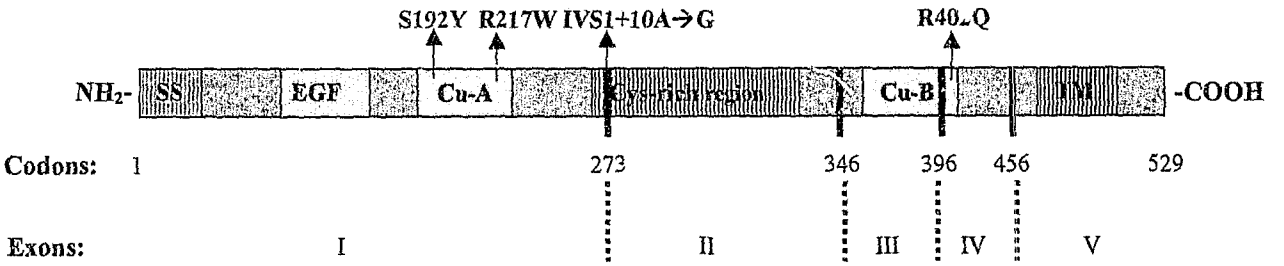


Fig.3.5: The position of the variants detected in the TYR protein in Caucasoid individuals. The codon positions of the start of every exon are indicated below the structure [Tomita, 1994; Sturm *et al*, 1995]. The exon positions are indicated below that. The OCA1 mutation is highlighted in red and the two polymorphisms are in blue. The intron variant is indicated in black, but has been taken to be a part of the consensus sequence in this study. The regions of the TYR protein as specified are: the N-terminal secretery signal peptide (SS), the epidermal growth factor-like region (EGF), copper binding regions A and B (Cu-A and Cu-B), Cystein-rich region, and the C-terminal membrane spanning domain (TM) [Sturm *et al*, 1995].

3.3 Variation at the *TYRP1* locus

The *TYRP1* locus was studied by investigating the frequency of the two ROCA mutations (368delA and S166X) in normally pigmented Negroid, Caucasoid and San individuals, as well as red-haired individuals. Following this, the entire gene was screened by SSCP analysis and sequencing in 15 Caucasoid and 15 Negroid individuals, to find out if other mutations were present that contributed towards the differences in pigmentation seen between them.

3.3.1 ASO results for 368delA detection

Exon 6 of the *TYRP1* gene was amplified by PCR and dotted in duplicate onto positively charged nylon membranes, which were then hybridised with oligonucleotides specific for the normal sequence or for the 368delA mutant sequence. The same blot was not used for the hybridisation of both probes to ensure that false positives did not occur. The procedure was carried out for a group of 24 random Negroid individuals, 15 San individuals and five individuals with red hair. Fig.3.6 presents examples of dot blots for the detection of both the normal and mutant sequences. All individuals tested were found to be negative for the 368delA mutation.

3.3.2 Restriction enzyme results for S166X detection

The S166X mutation was detected by restriction enzyme digestion. Exon 3 of the *TYRP1* gene was amplified by PCR and digested with *MboI*. The samples were electrophoresed in a 3% agarose gel. The samples included 48 unrelated Negroid individuals, 17 random San individuals, 47 random Caucasoid individuals and five Caucasoid individuals with red hair. Control individuals homozygous for the mutation, homozygous normal and heterozygous for the mutation were also digested and run on the gel. Fig.3.7 presents an agarose gel, which contains one individual from each of the four population groups studied and the control samples as indicated in the figure. The S166X mutation was not observed in any of the individuals studied.

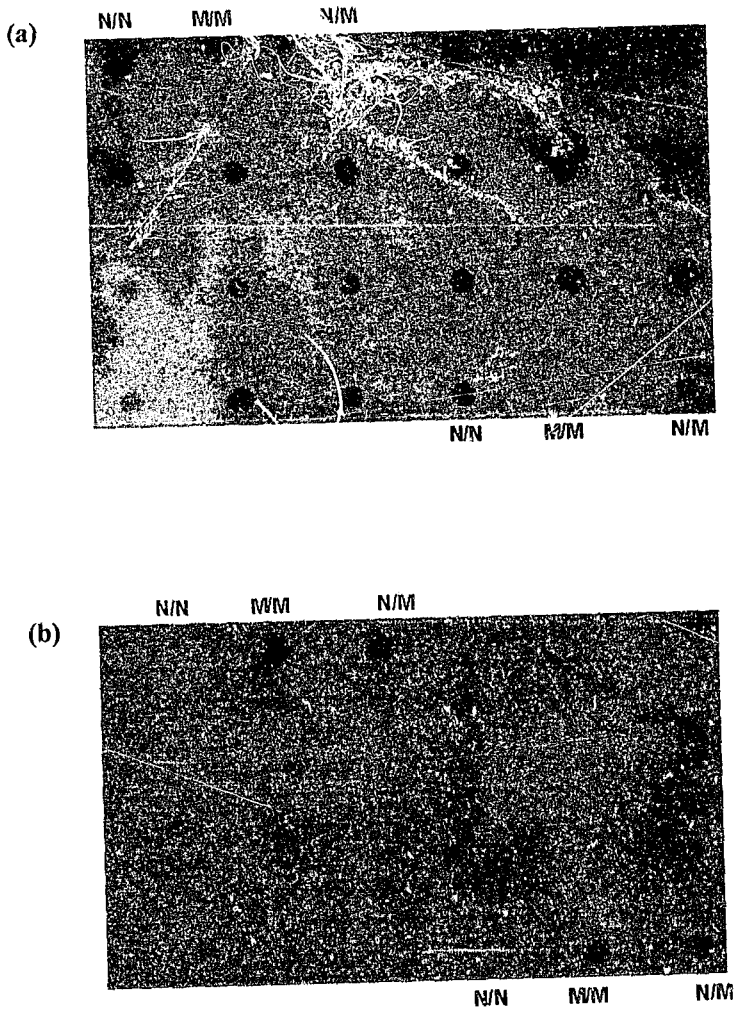


Fig.3.6: ASO analysis of the 368delA ROCA mutation. The dot blots are the results of hybridisation of DIG-labelled (a) normal and (b) mutant oligonucleotides. The samples used are a batch of 18 Negroid samples in addition to control samples homozygous normal (N/N), heterozygous for the mutation (M/N), and homozygous for the mutation (M/M).

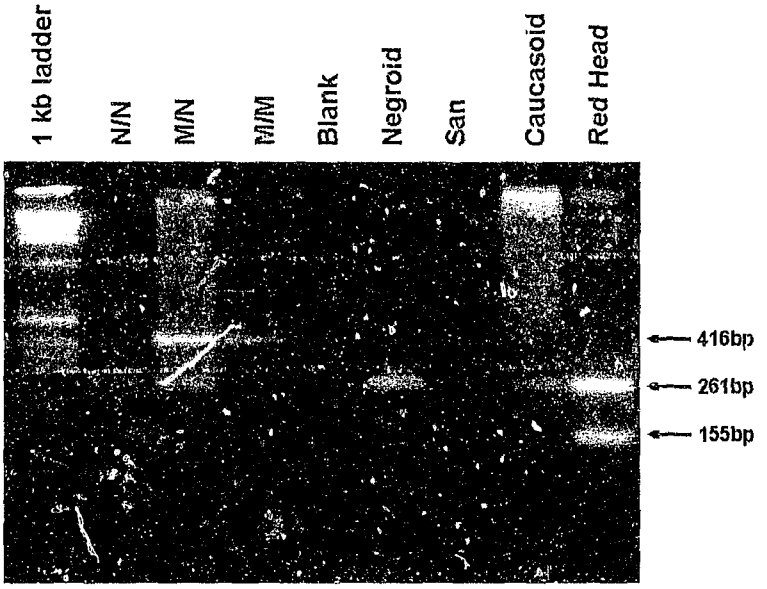


Fig.3.7: RFLP results of S166X detection. The 3% agarose gel above shows a set of control samples homozygous normal (N/N), heterozygous for the mutation (M/N) and homozygous for the mutation (M/M). Four samples, one from each population group studied, are shown, as well as a blank. The arrows on the right side indicate the fragment sizes that were observed. The 1kb ladder was essential to guage the size of the fragments that were produced.

3.3.3 SSCP analysis

As with the *TYR* gene, the *TYRP1* exons were amplified for SSCP analysis by PCR with ^{32}P-dCTP. The PCR product was electrophoresed on MDE gels containing 10% glycerol and gels without glycerol. Selected samples with band shifts on the SSCP gels were then sequenced. All exons of the *TYRP1* gene produced band shifts with no particular difference between the Negroid and the Caucasoid population groups. The most frequent variant was observed in exon 7 in four individuals of the Negroid group and two individuals in the Caucasoid group. The SSCP gel autoradiographs are not presented in this report, as they do not show much difference between the two groups. Exon 1 was not analysed because it is non-coding.

3.3.4 Sequencing results

Although the SSCP variants did not appear to indicate a difference between the Negroid and Caucasoid groups, one individual with each type of band shift observed was sequenced. The results are summarised in Table 3.7.

The sequences were compared to the consensus sequence from GenBank (accession number AF001295) (Appendix 4) [Box *et al*, 1998]. Two DNA sequence variants were observed that caused amino acid changes in the polypeptide. These are: R326H in exon 5 in one Negroid individual in heterozygous form and Y519X in exon 8 in two Negroid individuals in heterozygous form. Argenine (R) and histidine (H) are both basic amino acids, which have positive charges, and the variant may cause changes in structure and function of the TYRP1 protein because of amino acid size and shape, but not due to charge. Y519X results in a premature stop codon about 19 amino acids before the normal stop codon.

Two synonymous sequence variants were also noted: L7L in one Negroid individual and R87R in three Caucasoid individuals, all in heterozygous form. The R87R variant has been observed in Caucasoid individuals in an Australian study at a low frequency of 0.055 [Box *et al*, 1998]. In this study, the frequency of the R87R variant was found to be 0.1 (3/30) (Table 3.8), almost double the frequency in Box *et al*'s study. Their sample size was much larger (100 individuals), and may therefore present a more accurate frequency. The other two variants observed were in

the intron region, as shown in Table 3.7. One variant was common between the two populations, IVS7+43G→A, which was seen twice as frequently in the Negroid population. These mutations are illustrated in Fig.3.8.

3.3.5 Distribution of the variants in populations

Table 3.8 shows the frequencies of each variant and contains χ^2 values for each of the variants compared between the Negroid and Caucasoid populations. None was significantly different between the populations. As with *TYR*, the same tests were carried out here using the Chi-square and Arlequin programs. The phase of the variants was known because only one variant was observed in each individual, in heterozygous form. Thus, the information was presented as genotypic data with known phase in the data input file of Arlequin (Appendix 5).

The χ^2 calculations indicate that the *TYRP1* variants are not significantly different between the Negroid and Caucasoid populations ($p>0.05$). The population differentiation test by Arlequin confirmed the non-differentiation of these two populations at this locus, with a probability value of 0.156 (+/- 0.015).

The most frequent haplotype in each group was the normal consensus haplotype. It occurred at a frequency of 0.8 in the Caucasoids and 0.733 in the Negroids. The next most common haplotype in the Caucasoid population included the R87R variant, with a frequency of 0.1. In the Negroid population studied, the second most frequent haplotype (0.133) involved the IVS7+43G→A change in the intron.

Gene diversity was found to be 0.191 (+/- 0.093) in the Negroid and 0.00 in the Caucasoid as none of the variants in the Caucasoids affected the protein sequence. Hence, the gene is more diverse in the Negroid than the Caucasoid population, although the *TYRP1* locus as a whole is not very diverse in either population. Nucleotide diversity was 0.02% in the Negroids and 0.01% in the Caucasoids. Hence, the Negroids are more diverse at the nucleotide level than the Caucasoids.

The Negroid and Caucasoid populations studied were in Hardy-Weinberg equilibrium with p=1.000 in both cases.

3.4 *TYR* and *TYRP1* in the Negroid and Caucasoid samples

Table 3 9 has been presented to show the variants observed in each individual at both the *TYR* and *TYRP1* loci. It appears that overall there is more variation in the Caucasoid population group than the Negroid population group. The table clearly illustrates the difference in roles that the *TYR* and *TYRP1* loci play in the two populations. The Caucasoid individuals have much variation at the *TYR* locus compared to the Negroid individuals. The variation at the *TYRP1* locus is equally distributed between the two groups.

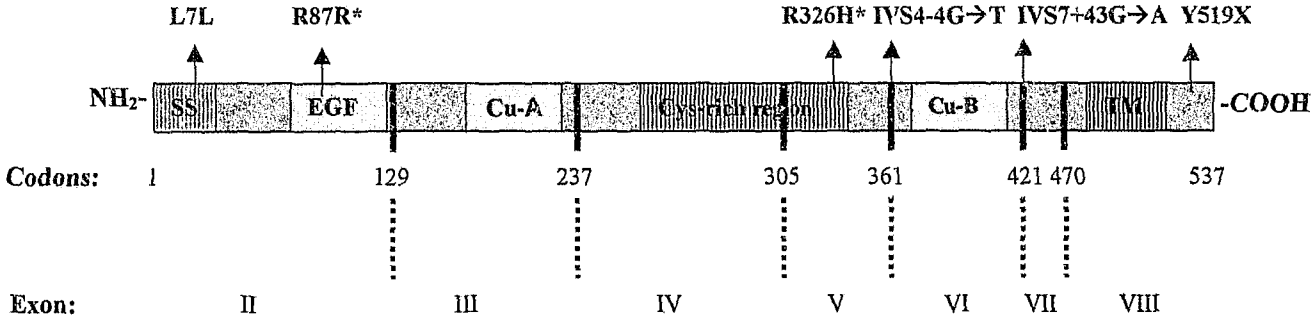


Fig.3.8: The position of the variants detected in the TYRP1 protein in Caucasoid and Negroid individuals. The codon position at the beginning of every exon is indicated below the structure. The exon positions are indicated below that. The variations in the coding region in Caucasoids are in blue, and those found in Negroid individuals are in green, while intronic variants are indicated in black. The regions of the TYRP1 protein as specified are: N-terminal secretery signal peptide (SS), the epidermal growth factor-like region (EGF), copper binding regions A and B (Cu-A and Cu-B), the cystein-rich region, and the C-terminal membrane spanning domain (TM) [Sturm *et al*, 1995]. The * indicates variants where the exact positions on the protein are unknown. Hence, an approximate location of the positions has been made.

Table 3.7: *TYRP1* gene variants found in random groups of Caucasoid and Negroid individuals

Exon	Base change	Position in coding sequence	Variant	Position in protein	Number of chromosomes		Reference
					Caucasoid	Negroid	
1	-	-	-	-	-		-
2	CTC→CTG	21	L7L	SS region	-	1/30	Present study
	CGG→AGG	259	R87R	Before or in EGF	3/30	-	Box *et al*, 1998
3	None	-	-	-	-		-
4	None	-	-	-	-		-
5	G→t 4th base before exon		IVS4-4G→T	In 3' acceptor site	1/30	-	Present study
	CGT→CAT	977	R326H	Region between Cu sites	-	1/30	Present study
6	None	-	-	-	-		-
7	G→a 43rd base after exon		IVS7+43G→A	Intron region	2/30	4/30	Present study
8	TAT→TAG	1557	Y519X	Past the TM region	-	2/30	Present study

Table 3.8: Variant frequencies and Chi-squared values for *TYRP1* between Caucasoid and Negroid individuals

Variants	Number of chromosomes		Variant frequencies		Chi-square calculations	
	Caucasoid	Negroid	Caucasoid	Negroid	Chi-square	Probability
L7L	0	1/30	0	0.033	1.017	0.313
R87R	3/30	0	0.1	0	3.158	0.076
IVS4-4G→T	1/30	0	0.033	0	1.017	0.313
R326H	0	1/30	0	0.033	1.017	0.313
IVS7+43G→A	2/30	4/30	0.067	0.133	0.741	0.389
Y519X	0	2/30	0	0.067	2.207	0.153

Table 3.9: *TYR* gene and *TYRP1* gene results for the random Negroid (N) and Caucasoid (C) individuals studied

Individual	TYR						TYRP1						
	1A	1B	2	3	4	5	2	3	4	5	6	7	8
N1													
N2													
N3													Y519X
N4													
N5												IVS7+43G→A	
N6												IVS7+43G→A	
N7										R326H			
N8													
N9							L7L						
N10													
N11												IVS7+43G→A	
N12												IVS7+43G→A	
N13													Y519X
N14													
N15					?								
C1		S192Y			R402Q								
C2													
C3					R402Q								
C4		S192Y			R402Q								
C5													
C6		S192Y								IVS4-4G→T			
C7		R217W											
C8		S192Y			R402Q		R87R						
C9		S192Y			R402Q								
C10		S192Y			R402Q								
C11		S192Y										IVS7+43G→A	
C12												IVS7+43G→A	
C13							R87R						
C14							R87R						
C15		S192Y			?								

- All the above variants occurred in heterozygous form and "?" indicates the individuals where results could not be obtained. This was included in the Arlequin input files.

4. DISCUSSION

Pigmentation is one of the most visible forms of variation in humans and has been studied for many years by different groups. Although some of the genetic mechanisms involved in this trait are known there is much that still remains a mystery. Hence, studying the different genes thought to be involved in this trait may clarify the role of these genes in normal pigmentation. Recent molecular studies have involved five genes that have a crucial role in pigmentation. They are the *MC1R*, *TYR*, *TYRP1*, *TYRP2* and *P* genes.

The main aim of this study was to investigate the role that the *MC1R*, *TYR* and *TYRP1* genes may play in normal pigment variation in humans. These genes were studied using a variety of molecular genetic techniques. Random Negroid and San groups and Caucasoid individuals with red hair were investigated in the *MC1R* study and random Negroids and Caucasoids were studied when investigating *TYR* and *TYRP1*.

4.1 *MC1R* gene variation in normal pigmentation

A study was carried out on the *MC1R* gene on a group of normally pigmented individuals to test for DNA sequence variation. *MC1R* was studied by investigating a group of 22 random Negroid individuals (59 random chromosomes deduced) and their mothers and a group of 17 random San individuals (34 random chromosomes), in order to assess the role that the gene may play in normal pigmentation. Eight Caucasoid individuals with red hair (14 random chromosomes) were also studied to confirm the role of *MC1R* in the red hair and pale skin phenotype in South African Caucasoid individuals, who originated from different parts of Europe and from Isreal.

Previous studies have shown that the variants of the *MC1R* gene in humans cluster in the first transmembrane domain, first intracellular loop, second transmembrane domain, the second intracellular loop, and the seventh transmembrane domain [Smith *et al*, 1998]. Some of the novel mutations found in this study are not confined to these areas. In addition to the latter hot spots, non-synonymous variants were found in the fourth, fifth, and the sixth transmembrane domains.

4.1.1 Variation in the *MC1R* gene in normally pigmented individuals

Previous studies have shown variation at the *MC1R* locus in various normally pigmented individuals, including Caucasoid, Indian and East and Southeast Asian individuals. The study of the Negroid and San populations revealed both synonymous and non-synonymous mutations, most of which have not been described previously. Overall, however, it must be noted that more synonymous mutations were observed than non-synonymous mutations. Previous studies have shown that synonymous base substitutions usually occur at a much faster rate than non-synonymous substitutions. The results obtained in the Negroid and San study of the *MC1R* gene also show this.

4.1.1.1 Synonymous and non-synonymous mutations

The rate of synonymous substitutions varies less between genes than rates of non-synonymous mutations. The overall rate of synonymous substitutions over all genes was found by Li (1997) to be 3.51×10^{-6} substitutions per synonymous site per year, while that of non-synonymous sites was found to be 0.74×10^{-9} substitutions per non-synonymous site per year. Note that these values were calculated for a total of 47 genes (such as genes for histones 3 and 4, insulin, albumin and so on) and are based on comparisons between human and rodent genes [Li, 1997]. The mouse and human *MC1R* sequences were compared by Rana *et al* (1999) by the calculation of the rates of synonymous (K_s) and non-synonymous (K_a) substitutions. They showed that $K_a:K_s < 1$ with $K_a = 0.155 +/- 0.0802$ and $K_s = 0.534 +/- 0.0802$. However, comparison of the non-synonymous mutation rates with the other members of the melanocortin receptor family revealed that the rate for *MC1R* mutations was the greatest, indicating that *MC1R* has evolved more rapidly than the other receptors in this family [Rana *et al*, 1999].

Although the present study has shown that there is far less non-synonymous variation at the *MC1R* locus than synonymous variation, a cumulative look at all studies that have been carried out on this gene imply otherwise. Tables 4.1 and 4.2 show the synonymous and non-synonymous mutations that have been found in the *MC1R* gene in this and other studies on normally pigmented individuals. These lists do not include Caucasoid individuals with red hair. There seems to be a greater number of non-synonymous mutations at the *MC1R* locus than synonymous mutations. Most of the synonymous mutations in Table 4.1 were observed in this

study in Negroid individuals and most non-synonymous mutations listed in Table 4.2 were observed in lighter pigmented individuals in previous studies. This observation may imply that non-synonymous mutations in the *MC1R* gene are associated with lighter pigmentation and that there is a lack of functionally significant variation in this gene in darker pigmented individuals. However, while attempting to link *MC1R* gene variation to melanoma susceptibility, Ichii-Jones *et al* (1998) found that variation at this locus did not determine skin type. It must be noted that this group investigated only three particular variants; D84E, V92M and D294H by RFLP analysis and may, therefore, have overlooked other variations in the *MC1R* gene that may have been present.

4.1.1.2 Non-synonymous mutations observed in this study and their implications

Changes to the DNA sequence may be considered to have some functional significance if they occur at evolutionary conserved regions of the DNA sequence. The alignment of the human MC1R protein with the mouse, dog, fox, sheep, horse and ox, showed that the amino acids at which the non-synonymous variants were observed in this study have been conserved between species. This suggests a possible functional significance for the S47I, L99I and F196L variants observed in this study. Rana *et al* (1999) has also compared the protein sequences of MC1R, including primate data, and their alignment results show that the positions at which amino acid changes were found in the present study have been conserved in primate sequences.

The F196L non-synonymous mutation that was observed in the Negroid group was of particular interest as it occurred at a relatively high frequency (7/59). In all but one individual it was found in the heterozygous state. It was present mainly in the South Sotho-speakers of the Negroid group studied, who are known to practice consanguineous marriages, which would not alter the gene frequency but may result in a large number of homozygous individuals. A study by Yang *et al* (1997) tested the effects of site-directed receptor mutagenesis on the binding affinity of NDP-MSH (a α-MSH analogue) to MC1R, as a previous study had suggested a model for the binding of these two proteins. This model suggested that the aromatic amino acid phenylalanine at position 196 was part of a series of phenylalanine residues in the protein that form a *hydrophobic ligand binding pocket*, which was hypothesised to be involved in the aromatic-aromatic binding of the hormone NDP-MSH to MC1R. Although the site-directed mutation F196A alone (both

residues not charged) did not change the binding potency of the hormone to the receptor, decreases in the binding affinity did occur when it was coupled with other mutations (namely Y179A and F182A). Therefore, it is not clear whether the mutation at this amino acid in the Negroid population may cause a change in the binding of the two proteins and in the melanogenic pathway. If mutations at this position of the *MC1R* gene did cause a drastic decrease in the binding of the hormone to the receptor, a decrease in tyrosinase stimulation may result, followed by a decrease in eumelanogenesis. If the individuals who had this variant were fairer in complexion to the rest of the group studied, it would confirm the importance of this site to α-MSH binding. However, it was not possible to obtain the pigmentation status of the individuals studied.

The S47I mutation was very rare (1/59), and was only observed in one individual from the group of random Negroids. The L99I variant was observed in one individual from the San (1/34).

The three non-synonymous mutations observed in the Negroid and San caused amino acid substitutions that produced no change in the charge of the residue at that position. Both residues of the L99I and F196L variants are non-polar (hydrophobic). S47I, on the other hand, produces a polar (hydrophilic) to non-polar change. The polar and non-polar residues play a significant role in the tertiary structure of the polypeptide chain. When the molecule folds it does this in such a way that the hydrophilic residues are on the surface of the molecule, where they react with water; while the hydrophobic residues are on the inside of the molecule, where they interact with each other. Hence, the S47I change may affect the folding of the protein and it's surrounding interactions. This variant only occurred in one individual and is, therefore, unlikely to play a major role in the Negroid population with regard to pigmentation. Functional studies are necessary to determine the effect of these non-synonymous substitutions on the function of MC1R.

4.1.1.3 Synonymous mutations observed in this study and their implications

The Negroid population group was found to have six silent mutations out of a total of eight mutations observed (Table 3.1). These mutations are thought to have no effect on the protein product because there is no change to the amino acid sequence of the polypeptide chain. The

alignment of the *MC1R* DNA sequences from different species allowed investigation of cross-species conservation of the positions at which the variants were observed. Most DNA bases at the mutation positions were conserved, except for positions 309, 699 and 942. A cytosine to thymine transition was observed in one child, which was inherited from the mother, at position 309. A guanine to adenine transition was observed in one Negroid individual at position 699. Likewise, at position 942 an adenine to guanine transition was observed in several Negroid and some San individuals. The nucleotides at these positions have not been conserved, and may therefore be considered to have an insignificant effect on the protein.

All synonymous mutations may, however, not be as neutral as would be expected and may actually produce changes in the protein and interfere with its function. Recent studies have shown that synonymous mutations can have an effect on the protein. For example, a silent mutation could create a donor splice site, which can cause a part of the coding region to be deleted [Li *et al*, 1995; Richard and Beckman, 1995]. [illegible] (1995) described a silent mutation in the *FGFR2* gene, associated with Crouzon syndrome, which created a donor splice site. Exon skipping can also result from changes in the consensus sequence at splice sites or at lariat branch-point regions [Liu *et al*, 1997]. Up to fifteen percent of the point mutations that cause genetic diseases in humans occur from abnormal splicing of mRNA. At least 101 different point mutations are know that have been found to affect mRNA splicing [Krawczak *et al*, 1992].

The silent mutations that were observed in the two African populations studied, therefore, may have an effect on gene splicing. The effects of the T314T variant on the mRNA would be an interesting subject to investigate because of its high frequency in the Negroid population (0.424). This mutation is found 10 bases upstream from the end of the coding sequence and may create a splice site in the coding region. The creation of splice sites by mutations have been observed in some genes, including the *F8* gene in haemophilia and the *HBB* gene in thalassaemia [Krawczak *et al*, 1992]. The functional significance of this and the other variants that were observed can only be determined by functional studies. RT-PCR of mRNA from these subjects, followed by sequencing would provide this information.

Table 4.1: Synonymous *MC1R* variants observed in normally pigmented individuals

Variant	Position on protein	Population (variant frequency)	Reference
L50L	First transmembrane domain	San (3/34)	Present study
A103A	First extracellular loop	Negroid (1/59)	Present study
L106L	First extracellular loop	Negroid (5/59)	Present study
I168I	Fourth transmembrane domain	Negroid (1/59)	Present study
Q233Q	Third intracellular loop	Negroid (1/59)	Present study
V265V	Sixth transmembrane domain	Negroid (1/59) San (2/34)	Present study
F300F	Seventh transmembrane domain	Negroids (2/59)	Present study
T314T	The COOH-terminal region	Caucasoids (present) Africans (21/50) East and Southeast Asians (28/120) Negroids (25/59) San (2/34)	Box *et al* (1997) Rana *et al* (1999) Present study

Table 4.2: The non-synonymous *MC1R* mutations observed in normally pigmented individuals

Variant	Position on protein	Population (variant frequency)	Reference
S47I**	First transmembrane domain	Negroid (1/47)	Present study
V60L*	First transmembrane domain	Caucasoids (16/178) Caucasoids (18/142) Caucasoid with red hair (1/14)	Box *et al* (1997) Smith *et al* (1998) Present study
R67V	First intracellular loop	Chinese (1/20) East and Southeast Asian (1/120)	Box *et al* (1997) Rana *et al* (1999)
D84E	Second transmembrane domain	Caucasoids (4/174) Caucasoids (present) Caucasoids (4/204) Caucasoids with red hair (1/4)	Box *et al* (1997) Koppula *et al* (1997) Smith *et al* (1998) Rana *et al* (1999)
V92M	Second transmembrane domain	Caucasoids (present) Caucasoids (17/154) Caucasoids (present) Caucasoids (1/8) East and Southeast Asian (1/120)	Valverde *et al* (1995) Box *et al* (1997) Koppula *et al* (1997) Rana *et al* (1999)
V92L	Second transmembrane domain	Caucasoids (3/174)	Box *et al* (1997)
L99I**	Second transmembrane domain	San (1/34)	Present study
I155T	Second intracellular loop	Caucasoids with red hair (1/174) Caucasoids with dark hair (1/142)	Box *et al* (1997) Smith *et al* (1998)
R163Q	Second intracellular loop	Chinese (16/20) East and Southeast Asians (84/120) Indians (present) Caucasoids (4/142)	Box *et al* (1997) Rana *et al* (1999) Smith *et al* (1998)
F196L**	Fifth transmembrane domain	Negroids (7/59)	Present study

*Variant observed in the present and previous studies.
**Variants observed in the present study alone.

4.1.1.4 Comparison of the Negroid and San population groups at the *MC1R* locus

The results from this study suggest that pigmentation in the Negroid and San populations is not influenced in a major way by mutations at the *MC1R* locus, as most of the observed variants were synonymous. Southern Africa was occupied by the Khoisan before the appearance of the Bantu-speakers approximately 2 000 years ago [reviewed in Soodyall and Jenkins, 1998]. Contact occurred between the Khoisan and the Negroid during the process of the migration of the latter group into central and southern Africa. This has lead to admixture between them [Soodyall and Jenkins, 1998] and may explain the common variants observed in the two groups, namely V265V and T314T. However, mtDNA studies by Soodyall and Jenkins (1998) have found no shared mtDNA types in the Khoisan and the Negroid populations studied, although a few mtDNA types may share a recent common ancestry.

The most frequent variant in the Negroid individuals, T314T (also defined as A942G), has been reported by Rana *et al* (1999) in a group of African individuals (Table 4.1). In their study they suggested that the sequence with 942G is probably the ancestral nucleotide because the chimpanzee and gorilla had a G at that site. Unexpectedly, however, this variant was rare in the San population (0.059) compared to the Negroid population (0.424). A possible explanation is random drift in a small geographically isolated population. This variant has also been reported in some Caucasoid, East and Southeast Asian and Indian individuals [Box *et al*, 1997; Rana *et al*, 1999], but at much lower frequencies than in the Negroid.

The statistical analyses show that the Negroid and San population groups studied are very different at the *MC1R* locus. The most frequent variant in the Negroid was T314T, with a frequency of 0.424, which is comparable to the frequency observed by Rana *et al* (1999). It must be noted, however, that Rana's group of African individuals included various groups of African origin, including the pygmy and the San. The most frequent variant in the San was the L50L variant, with a frequency of 0.088. χ^2 analysis showed that the T314T variant is significantly associated with the Negroid population ($p<0.001$), as is the F196L variant (particularly in the South Sotho-speakers) with $p<0.05$. L50L is significantly associated with the San population ($p<0.05$).

Nucleotide diversity calculations suggest that the Negroid population group (0.12%) is older than the San population group (0.05%) because of the greater diversity in the Negroids. However, it is known that the San population is less diverse genetically because of their small population size and relative geographic isolation. The low diversity has been observed by other genetic studies [reviewed in Nurse *et al*, 1985; Ramsay and Jenkins, 1985].

4.1.1.5 Variation at the *MC1R* locus in individuals from Asia and Europe

The *MC1R* locus has been investigated, by Box *et al* (1997) and Rana *et al* (1999), in normally pigmented Caucasoid and Asian individuals who can be classed as individuals with light to intermediate skin pigmentation. These individuals were not studied in this project. Table 4.2 shows that the only alleles shared by the Caucasoid and Asian individuals are the R163Q and V92M variants. The V92M variant has been found to produce a receptor that allowed for a decreased potency of α-MSH hormone action [Xu *et al*, 1996]. As this and the R163Q mutation is not present in the darker pigmented Negroid individuals, they may be associated with lighter pigmentation. The Indian individuals with these mutations may have had a lighter complexion than those without, although this information was not provided by Rana *et al* (1999). It appears that the majority of the non-synonymous mutations are associated with lighter pigmentation, and the synonymous mutations are associated with darker pigmentation.

4.1.2 *MC1R* gene variation in Caucasoid individuals with red hair

Studies by Valverde *et al* (1995), Box *et al* (1997), Smith *et al* (1998) and Rana *et al* (1999) have shown that mutations at the *MC1R* locus do have a strong association with the red hair and pale skin phenotype. The present project has confirmed these observations and mutations in the *MC1R* gene have been detected in all eight red-haired individuals, with the observation of eight mutations at this locus, four of which are novel: S83P, Y152X, A171N, and P256S.

The mutations observed were compared to *MC1R* sequences from other species by alignment of DNA and protein sequences. Most of the non-synonymous mutations found by this study in Caucasoid individuals with red hair were in conserved regions between the human, mouse, dog, fox, sheep, horse and ox at the DNA and amino acid levels. At the protein level position 171 was

not conserved, with the mouse and horse having valine at the position, instead of alanine. The primate amino acid alignment from Rana *et al* (1999) also shows that the positions at which amino acid changes were found have been conserved in primate sequences. These mutations are therefore possibly functionally significant.

4.1.2.1 Functional significance of the variants in Caucasoids with red hair

The novel variant Y152X and the previously described variants, R151C and R142H, occur in the second intracellular region of the protein. This region of the protein has been found to contain two sequences for cAMP-dependent protein kinase recognition: from amino acid 142 to 145 and from 151 to156. Additionally, one other such region is located in the COOH-terminal region (amino acids 306 to 308) [Chhajlani and Wikberg, 1992]. The three variants mentioned above are in the critical region and may lead to the production of a protein without sufficient function for signalling cAMP. This may also apply to R160W, as it lies very close to the second intracellular region. Functional studies on the R151C mutation by Frandberg *et al* (1998), where receptor binding and cAMP assays were carried out, revealed that although α-MSH (radio-labelled) bound to the receptor with the R151C mutation with identical affinity as the wild-type MC1R it could not stimulate the production of cAMP. Hence, this mutation causes the MC1R protein to be non-functional. The result is the absence of tyrosinase activation and phaeomelanogenesis [Frandberg *et al*, 1998]. Likewise, functional studies need to be carried out on the other variants in this region, observed in this and other studies. The other variants seen in the red-haired individuals were scattered throughout the protein in transmembrane domains 1, 2, 4, 6, and 7 (Fig.3.2).

The variants that were found are illustrated in Fig.3.2. V60L and R142H do not result in changes in charge or polarity. The former has been recognised by Box *et al* (1997) to be present in individuals that had blond and fair/brown hair, and the R142H variant was identified by the same group in one red-haired individual. The novel variants S83P, P256S and A171N cause a polar to non-polar change and non-polar to polar changes respectively, while R151C and R160W cause a positive charged to an uncharged polar residue. D294H causes a negative to positive charge change. The three that cause polar to non-polar changes (or vice versa) would affect the tertiary structure of the MC1R polypeptide because of the effect on the hydrophilicity. The latter three

variants may also have an effect on the structure of the protein. They may affect the secondary folding or the binding of α-MSH.

4.1.2.2 Statistical implications

The group of red-haired individuals was very different at this locus to the San and Negroids. However, it must be noted that the group of red-haired individuals was not a random one, as they were selected for that phenotype. Gene diversity and nucleotide diversity of the group of red-haired individuals in this study was greater than the Negroid and San at the *MC1R* locus.

In the Caucasoid individuals with red hair the R151C and R160W had the highest frequencies, with all other variants occurring only once. As normally pigmented Caucasoid individuals were not included in this study, the red-haired individuals were not analysed by χ^2 analysis. The Negroid and San and the red-haired individuals did not share any of the variant alleles. This, along with the fact that the red-haired individuals probably had functionally significant mutations at the *MC1R* locus, further emphasises the probability that variation at the *MC1R* locus is associated with the red hair and pale skin phenotype.

4.1.2.3 Is variation at the *MC1R* locus responsible for the red hair and pale skin phenotype?

The question has arisen about which specific variants are associated with the red hair and pale skin phenotype. Some variants, such as V92M and V60L, were thought to be linked to red hair and pale skin [Valverde *et al*, 1995], and were later found in individuals with other types of normal pigmentation [Box *et al*, 1997; Koppula *et al*, 1997]. R151C, R160W and D294H have been linked to red hair and pale skin by most of the previous studies. Smith *et al* (1998) found that these variants are significantly associated with red hair ($p=0.0015$, $p<0.001$ and $p<0.005$ for R151C, R160W and D294H respectively). In their study all individuals who were compound heterozygotes for two of these three variants had red hair. In the present study one red-haired individual was a compound heterozygote for R160W and D294H, one was homozygous for R151C and another homozygous for R160W. The results of this study, therefore, are in agreement with the fact that these three variants are associated with red hair and pale skin.

Box *et al* (1997), however, presented one individual without red hair who was homozygous for R151C. Their twin study has suggested that *MC1R* may not be the only locus involved in the red hair and pale skin phenotype, because dizygotic twins discordant for red hair had the same *MC1R* variants. *MC1R* mutations appear to be necessary but not sufficient to cause the red-haired phenotype [Box *et al*, 1997]. In addition to this, although Smith *et al* (1997) found that one *MC1R* variant allele was observed, in heterozygous form, in normally pigmented individuals without red hair and pale skin, some individuals with red hair also had only one variant allele. They, therefore, suggested an altered expression of the wild type-allele by another mechanism. Hence, the inheritance of the red hair and pale skin phenotype is not the simple autosomal recessive trait, as might be suggested by the results of compound heterozygosity and homozygosity of variants from the present study. The switch from eumelanogenesis to phaeomelanogenesis is still an enigma and a clearer understanding of this switch may help us understand the mechanism behind the occurrence of the red hair and pale skin phenotype better.

4.1.3 Conclusions from *MC1R* data

The Negroid and San show some significant differences at the *MC1R* locus (Table 3.2). The San have relatively fewer variants than the Negroids, which can be explained by the small population size and the geographic isolation of the San population. They share two of the variants either as a result of recent admixture or because of common ancestry. The results found here suggest that variation at the *MC1R* locus does not play a major role in the dark pigmentation observed in the Negroid as most variants found in this gene were synonymous. The San's yellow-brown pigmentation is unlikely to be due to variation at the *MC1R* locus.

The *MC1R* gene clearly plays a significant role in the red hair and pale skin phenotype of South African Caucasoid individuals with European ancestry, as mutations at this locus were found in all eight individuals with this pigmentation type. Although the results from this study suggest an autosomal recessive trait, other studies suggest that variants in the coding region of this gene are not the only cause of red hair and pale skin [Valverde *et al*, 1995; Box *et al*, 1997; Smith *et al*, 1998; Rana *et al*, 1999]. Since parents of the red-haired individuals were not studied it was not possible to determine conclusively that each gene had a mutation. Some of the mutations may

have been present in cis. Family studies would have helped to discover the phase of the mutations. The functional significance of the variants observed in this study need to be determined by functional studies. As pale skin has a high risk of early skin cancer development, the discovery and investigation of *MC1R* variants may assist in understanding the mechanism and in developing suitable methods of prevention of this disorder.

Although this study provided some interesting points for discussion, it should be considered as a pilot study. A control set of normally pigmented South African Caucasian individuals was not tested to compare with the red-haired individuals. A comparison of the two groups would have provided a more complete picture of the significance of the role that the *MC1R* gene plays in the red hair and pale skin phenotype in Caucasoid individuals in South Africa. Additionally, the red-haired individuals should have been subjected to a more rigorous phenotypic or clinical screening method, with respect to their skin type. Skin reflectance tests and the analysis of skin type using Fitzpatrick's scale of skin typing would have been beneficial. This would have helped to identify which mutations segregated with which skin type, and may have provided a much clearer view of the function of the MC1R protein.

4.2 *TYR* and *TYRP1* gene variation in normal pigmentation

The *TYR* and *TYRP1* genes were investigated to search for functionally significant alleles that may play a role in normal pigment variation. Hence, a group of 15 random Negroid individuals and 15 random Caucasoid individuals were studied. These groups represent the two extremes of normal pigmentation variation of the different populations in the world, and would help to determine the molecular mechanism behind normal pigment variation in humans.

4.2.1 The role of the *TYR* gene in normal pigment variation

The *TYR* gene was studied by SSCP analysis and sequencing. Two common non-synonymous (S192Y and R402Q) variants were identified in the *TYR* gene of some of the Caucasoids that have been described previously. In the present study the S192Y variant occurred at a lower

frequency (0.267) in the Caucasoid population in this study than the frequency in all populations (except the Orientals) (0.48) reported by Spritz and Hearing (1995). The R402Q variant was more frequent in the Caucasoid population (0.214) in this study than in all populations (except the Orientals) (0.15) listed by Spritz and Hearing (1995) (Table 4.3). In addition to these, a very rare OCA1 mutation, R217W, was detected in one individual in heterozygous form. The Negroid individuals studied all had the consensus *TYR* gene sequence.

4.2.1.1 Effect of mutations on the protein

The three mutations observed in the Caucasoid group were in the copper binding regions (Fig.3.6), S192Y and R217W in copper binding region A and R402Q in copper binding region B. Tyrosinase catalyses three different reactions by acting as a tyrosine hydroxylase, DOPA oxidase and DHI oxidase. R402Q has been found to cause a reduction in the DOPA oxidase and DHI oxidase activities of tyrosinase at 31°C. In this mutant these activities are absent at 37°C [Tripathi *et al,* 1992a]. This may be one mechanism involved in causing the lighter pigmented appearance of the Caucasoid individual. Although, Tripathi *et al* (1992a) suggest that the latter observations imply that the oxidase activities occur at a site, Oetting and King (1994) suggest that all three catalytic activities of tyrosinase occur at one single catalytic site. The R402Q change causes a positive to no charge change at the position, which may cause a change in the folding of the polypeptide chain.

The amino acids serine and tyrosine in the S192Y variant are non-polar and occurs in the one copper binding domain of tyrosinase, which binds to copper atoms in order to carry out its catalytic functions. The hydroxyl groups of both serine and tyrosine in the copper binding domains may act to help bind the copper onto tyrosinase [Furumura *et al,* 1998b]. Therefore, the change from serine to tyrosine may not have an effect in the binding of copper to the enzyme at the copper A domain. However, the variant may have another effect on the function of the protein due to the change in size and chemical structure of the amino acid. This variant can be detected by *MboI* cleavage [Fukai *et al,* 1995].

Several of the Caucasoid individuals (5/15) in this study were double heterozygotes for the S192Y and R402Q variants. It was not ascertained whether these occurred in cis or trans. Other

studies have shown that while the tyrosinase protein with Y192/R402 is thermostable, the Y192/Q402 protein has only 27% catalytic activity at 37°C [Tripathi *et al*, 1991, cited in Fukai *et al*, 1995]. These data may further suggest that the S192Y variant does not affect the function of the TYR protein, as the changes in the potency of the enzyme seem to be due to the R402Q variation.

R217W is a very rare OCA1A mutation that was found in one individual in the Caucasoid group studied, in heterozygous form. OCA1 mutations cluster in four main regions of tyrosinase: the two copper binding sites A and B, the 5' end of the coding region and the 3' end of copper binding site B [Oetting and King, 1994]. The R217W mutation occurs in copper binding site A. The change from positive to an uncharged non-polar amino acid and the change in size and shape may affect the binding of copper, causing a reduction in the activity of tyrosinase.

Comparison of the TYR, TYRP1 and TYRP2 amino acid sequences by alignment [Sturm *et al*, 1995] show that the 402 position (mutation R402Q) of TYR has been conserved in this family of protein and the 217 position (mutation R217W) has been conserved in TYR and TYRP1, but not in TYRP2. Mutations at a conserved position may have an effect on the function of the protein, suggesting that these two mutations observed may have a direct effect on protein function. The 192 position (mutation S192Y) has not been conserved, and the variation observed may not be of functional significance to the protein, as has already been suggested.

4.2.1.2 Statistical analyses

χ^2 analysis and the exact test of population differentiation have shown that the Negroid and the Caucasoid populations are very different at the *TYR* locus ($p<0.05$). The R402Q and S192Y variants were found at relatively high frequencies in the Caucasoid group of individuals, but were absent in Negroids.

Gene diversity calculations show that the Caucasoid individuals are more diverse at the *TYR* locus (0.582) than the Negroid group of individuals (0.00). Nucleotide diversity is also higher in the Caucasoids (0.05%) than the Negroids (0.00%). Showing that this locus may play an important role in the differences observed in these two population groups.

The Hardy-Weinberg equilibrium studies show that the Caucasoid population sample was in equilibrium, at this locus, indicating that there was no sample bias. The Negroid group studied here had the consensus sequence of the *TYR* gene in all cases. This may be the result of the small sample size (15 individuals) and the fact that other alleles that exist in the South African Negroid population occur at very low frequencies.

4.2.2 *TYRP1* gene variation in normal pigmentation

The role of *TYRP1* in normal pigmentation was investigated by two studies. One study was to determine the frequency of the ROCA mutations, S166X and 368delA, in normally pigmented individuals. In the other, the *TYRP1* gene was studied by SSCP analysis and sequencing to assess the role of this gene in normal pigment variation. The study did not show a distinct difference between the normally pigmented Negroid and Caucasoid groups at the *TYRP1* locus. The variants that were found were rare (Table 3.8).

4.2.2.1 The ROCA mutation study

The ROCA mutations, S166X and 368delA, both result in the truncation of the *TYRP1* protein. These mutations accounted for 95% of the ROCA mutations in a study by Manga *et al* (1997), and were investigated in the present study to find the frequency of these mutations in normally pigmented individuals. The San and red-haired Caucasoid individuals were also investigated to find out if these mutations played a part in the yellow and red pigmentation in these individuals. The S166X and 368delA mutations were not found in any of the individuals investigated. This shows how rare these mutations are in normally pigmented individuals and that they are specific to the ROCA hypopigmentation phenotype. Note that the random Caucasoids were not tested for the 368delA mutation, but only for S166X.

4.2.2.2 *TYRP1* gene screen of Negroid and Caucasoid individuals

The present study revealed various mutations, at low frequencies, both in the coding regions and the introns. Two synonymous variants, two non-synonymous variants and two intron mutations were found.

The R87R variant, caused by a transversion from cytosine to adenine, was observed in three of the Caucasoid individuals in the present study, with a frequency of 0.1 (almost double the frequency observed by Box *et al*, 1998). Box *et al* (1998) studied Caucasoid individuals and found the R87R synonymous variant at a low frequency of 0.055 (Table 4.4). The study by Box and colleagues consisted of 100 individuals, and may represent a more accurate estimation of the frequency of the allele, although there may be a real difference between the groups. One Negroid individual also had a synonymous variation L7L, in the secretery signal peptide, which is removed when the protein is processed [Tomita, 1994].

The intron variation observed upstream from the 5' end of exon 5 was found to be within the 3' splice acceptor site of the intron, and may have been of significance to the Negroid pigmentation type had it occurred more frequently. This variant may have affected exon splicing. IVS7+43G→A variant was observed in both the Negroids and Caucasoids at frequencies of 0.133 and 0.067 respectively. It is present in the intron region, away from both the splice acceptor and donor sites and may not affect the protein. As this variant was observed at low frequency in both population groups it in unlikely contribute to the pigment differences observed, between them.

Two non-synonymous variants, R326H and Y519X, were observed at low frequencies in the Negroid group. The alignment of the TYRP family of proteins [Sturm *et al*, 1995] shows that position 326 has been conserved between the *TYR* and *TYRP1* genes, and the variation observed in this study may be of functional significance. The Y519X is a protein truncation mutation that is likely to affect the function. The variants occurred at low frequencies (0.033 and 0.067 respectively) (Table 4.4), and are therefore unlikely to make a major contribution to pigment variation in Negroids.

It appears from this study that variations in the *TYRP1* gene do not contribute in a major way to the pigment differences between the Negroid and Caucasoid individuals, as variants were found in both groups at low frequencies. This study has, however, revealed two non-synonymous mutations and one synonymous variant in the *TYRP1* gene of Negroid individuals, which have not been reported previously.

4.2.2.3 Statistical analyses

Statistical analysis of the *TYRP1* locus shows that the Negroid and Caucasoid populations studied are not different at this locus. This is based on the χ^2 analysis results and the results of the exact test of population differentiation, which show that the two populations are not significantly different from each other at this locus ($p>0.05$ in all cases). The variants observed were not very frequent in either population, which suggests that they are unlikely to contribute in a major way to the pigment variation in these two populations. The Hardy-Weinberg equilibrium calculations show that both populations were in equilibrium.

4.2.3 Implications of *TYR* and *TYRP1* variation

The tyrosinase gene (*TYR*) sequence is very different between the Negroid and Caucasoid populations in South Africa. Two relatively common variants were found, that appear to segregate with the Caucasoid pigmentation type, and no variation was found in the Negroids. Hence, the *TYR* gene appears to play a role in the pigmentation differences between the Caucasoid and Negroid populations in South Africa, and variation at this locus likely contributes to the lighter pigmentation of the Caucasoids. However, as some individuals in the Caucasoid group did not have either one of these variants, another locus or other loci, as well as environmental factors play a role.

The *TYRP1* gene, however, does not appear to play a significant role in normal pigment variation as little variation in the gene sequence was noticed in and between the Caucasoid and Negroid groups studied. All variants that were observed occurred at very low frequencies in both population groups. The function of the *TYRP1* gene is not yet known conclusively in humans, although it has been proposed to act as a DHICA oxidase. The ROCA phenotype of hypopigmentation and a reddish tinge to the skin and hair is caused by mutations at the *TYRP1* locus in and hence this gene does play a role in pigmentation. However, it does not appear to play as significant a role as *TYR* seems to play in normal pigment variation.

Table 4.3: Synonymous variants observed in normally pigmented individuals in the *TYR* gene

Variant	Position on protein	Population (variant frequencies)	References
S192Y	Cu-A binding site	All populations (0.48) Orientals (0.00) Caucasoids (0.27)	Giebel and Spritz (1990) Present study
R402Q	Cu-B binding site	All populations (0.15) Orientals (0.00) Caucasoids (0.21)	Tripathy et al (1992) Present study

Table 4.4: Synonymous and non-synonymous variants observed in normally pigmented individuals in the *TYRP1* gene

Variant	Position on protein	Population (variant frequencies)	References
L7L	Secretary signal peptide	Negroid (0.033)	Present study
R87R	Before or in Epidermal growth factor	Caucasoids (0.055) Caucasoids (0.1)	Box et al (1998) Present study
R326H	Region between Cu sites	Negroid (0.033)	Present study
Y519X	Past the transmembrane region	Negroid (0.067)	Present study

4.3 Conclusions: Genes in normal pigmentation

This study aimed to explore the role of the *MC1R*, *TYR* and *TYRP1* genes in normal pigment variation. Pigmentation is a complex polygenic, multifactorial trait that has been studied by various groups in an attempt to uncover the molecular mechanism that gives rise to the diversity that has been observed in the many populations in the world in health and disease. Although much needs to be done to realise the importance of each gene in normal pigmentation, the findings of this study add to the body of information that is being built on pigmentation in South Africa and the world. The findings suggest that functionally significant mutations in the *TYR* and *MC1R* loci may well contribute to lighter skin pigmentation and that the *TYRP1* gene plays a lesser role. The greater diversity in Caucasoids compared to Negroids in the *TYR* and *MC1R* genes emphasises the significance of these genes in normal pigment variation in light of the fact that the latter is older in evolutionary terms and therefore is expected to show larger nucleotide diversity.

The *MC1R* and *TYR* studies show an overall greater variation in the lighter pigmented population group than the darker pigmented group. This may be correlated with the vitamin D hypothesis, which suggests that as the populations migrated from Africa to the Northern Hemisphere (not more than 100 to 200 thousand years ago, as hypothesised by the "out of Africa" model of evolution) the skin had to adapt to the relative decrease in UV radiation to produce necessary quantities of vitamin D within the skin. As the darker individuals moved from the Southern Hemisphere to the Northern Hemisphere they accumulated more mutations in the genes involved in pigmentation in order to be able to synthesise the required amount of vitamin D. These mutations were tolerated with respect to pigmentation, because the decrease in amount of radiation encountered as the populations moved north obviated the need for darker pigmentation.

Further studies on normal pigment variation should include the study of the *P* and *TYRP2* genes, with the same 15 normally pigmented random Negroids and Caucasoids, to ascertain the role that these genes play in these two extremes of pigmentation. *α-MSH* and *ASP* are other genes that could be studied, especially with respect to the red hair and pale skin phenotype since mutations in *α-MSH* were found to result in the red hair phenotype [Krude *et* 1998]. The function of ASP in humans is not know, but if it does have the same role as in the mouse as an antagonist for α-MSH binding, this gene may also play a significant role in

the red hair and pale skin phenotype. The *P* gene is another interesting contender for mutations that may lead to a lighter phenotype, as its protein is thought to play a role in the stabilisation of the melanogenic complex. *P* mRNA has been found to be absent in phaeomelanic tissue, implying that P plays a critical role in eumelanogenesis [Lamoreux *et al*, 1995]. In the absence of the functional mouse p protein the pH of the melanosome has been found not to be acidic and a low pH is necessary for tyrosinase function. This is more evidence to favour the importance of the P protein in eumelanogenesis, which is favoured by high tyrosinase levels [Puri and Brilliant, 1998]. Hence, this gene may very well play a significant role in the pigment phenotype.

As pigmentation is a polygenic trait, family studies could be carried out with careful correlation between the phenotype and mutations at a wide variety of known and candidate pigment loci. The sample size of 15 individuals, used in the *TYR* and *TYRP1* studies, might not be considered a large enough sample size for a fair assessment of the genetic variability of these genes in the populations studied, even for a preliminary study such as this. A larger sample size of 35 to 50 individuals may provide a better idea of the genetic variability in a population. Larger sample sizes of populations, together with a quantitative measurement of pigmentation would be useful to find a correlation between pigmentation and the mutations observed.

Much could be done to finally understand the role that each gene plays in normal pigment variation in humans. In future, the findings may provide some clues to treatment and prevention of skin cancer, as well as treatment for various hypopigmentary disorders.

APPENDICES

APPENDIX 1: Ethics clearance and information sheets

The information sheet for the *MC1R* gene red hair and pale skin study is presented below. The ethics clearance certificates are attached overleaf.

INFORMATION SHEET

DNA SEQUENCE VARIATION IN PIGMENT GENES

The pigment biosynthetic pathway is very complex and is only partly understood. There are essentially two types of pigment, phaeomelanin and eumelanin and these in different proportions of packaging within the cells, give rise to the great diversity in human skin, hair and eye pigmentation. A few genes which are important in the synthesis of pigment have been identified and can now be studied at the DNA level.

The aim of this study is to try to determine some of the causes of variation in skin, hair and eye colour by looking at the DNA sequences of the pigment genes in different individuals. This information should lead to a better understanding of the formation of pigment and may provide clues to skin cancer susceptibility and for some medical conditions which are characterised by pigment abnormalities.

The study will involve taking a blood sample, a small sample of hair for comparative purposes as well as a photograph. We would like to request this information in Scientific or Medical publications, should it be relevant. (Questionnaire / consent form separate).

Participation is entirely voluntary and you will not be disadvantaged in any way if you wish not to participate. There will be no cost implications to you if you do not participate.

For any further information contact Professor Michele Ramsay, Department of Human Genetics. (Tel. 011 489 9214).

STUDY ON SKIN AND HAIR PIGMENTATION

CODE:
DATE:

NAME:

DATE OF BIRTH:

ADDRESS:

TELEPHONE NUMBER:

Hair colour:

Skin colour:

Presence of freckles:

Tanning ability:

Sunburn predisposition:

FATHER: Origin -

Hair colour -

MOTHER: Origin -

Hair colour -

Blood samples:	YES / NO
Hair sample:	YES / NO
Photograph:	YES / NO

FAMILY PEDIGREE: Please draw on back of sheet.
(Include parents, grand parents and siblings of index case and give hair colour of each as it was during their youth)

I have been informed of the purpose of the study and agree that my DNA may be used for the study of pigment variation and for general population studies. I shall/shall not be agreeable to my photograph and photograph of my hair sample being used for publication purposes.

Name:..... Date:.....

Witness:.....

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

COMMITTEE FOR RESEARCH ON HUMAN SUBJECTS (MEDICAL)
Ref: R14/49 John

CLEARANCE CERTIFICATE PROTOCOL NUMBER M 970510

PROJECT DNA sequence variation in pigment genes in health and disease

INVESTIGATORS Miss MPR John

DEPARTMENT Human Genetics, Medical School

DATE CONSIDERED 970530

DECISION OF THE COMMITTEE

Approved unconditionally

DATE 970613

CHAIRMAN. ..... (Professor P E Cleaton-Jones)

c c Supervisor: Professor M Ramsay
Dept of Human Genetics, SAIMR

==========

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and ONE COPY returned to the Secretary at Room 10001, 10th Floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee.

DATE 18-6-97 SIGNATURE M. Ramsay .....

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

COMMITTEE FOR RESEARCH ON HUMAN SUBJECTS (MEDICAL)
Ref: R14/49 Ramsay

CLEARANCE CERTIFICATE PROTOCOL NUMBER M 940711

PROJECT Detection of mutations within the genes for brown, rufous & tyrosinase-positive oculoutaneous albinism & structural ....

INVESTIGATORS Dr M Ramsay

DEPARTMENT Human Genetics, SAIMR

DATE CONSIDERED 940729

DECISION OF THE COMMITTEE *

Unconditionally approved

DATE 940818

CHAIRMAN. ..... P Bloom (Professor P E Cleaton-Jones)

* Guidelines for written "informed consent" attached where applicable.

c c Supervisor: Professor T Jenkins
Dept of Human Genetics, SAIMR

Works2\other\heclear\ 940711

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and ONE COPY returned to the Secretary at Room 10001, 10th Floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee.

DATE.....SIGNATURE ..... M. Ramsay .....

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

APPENDIX 2: Solutions

Solutions	Recipe
Agarose gel (1-3%)	1to 3g HGT agarose in 100ml 1XTBE buffer and melt in a microwave oven Add 3 to 9µl 1% ethidium bromide and mix Leave to cool and pour into horizontal plate and allow to set
Denaturing solution for dot blots	10mM Tris (121.1mg) 0.3M NaOH (1.2g) Make up to 100ml solution with distilled water
Dextran-formamide dye	20mg dextran Make up to 1ml in formamide (work in the fume hood)
dNTP (125mM)	125µl dATP, dCTP, dGTP, dTTP (10mM stock) Make up to 1ml double distilled water
EDTA (0.5M)	93.06g EDTA disodium salt in distilled water pH 8 with 5M NaOH Make up to 500ml
Ethidium bromide	1g EtBr Make up to 100ml with distilled water and store in a dark bottle
Ficoll dye	50% sucrose 50mM EDTA (pH 7.0) 0.1% bromophenol blue 10% ficoll dye
Formamide/EDTA/BPB loading dye	10ml formamide 10mg xylene cyanol 10mg bromophenol blue 0.1mM EDTA (200µl of 0.5M stock)
1kb Ladder	2.1ml 1XTE 250µl 1kb ladder 125µl ficoll dye

MDE gel (- glycerol)	4% MDE (25ml) 6ml 10XTBE Make up to 100ml in distilled water Degas and store at 4°C
MDE gel (+ glycerol)	10% glycerol (10ml) 4% MDE (25ml) 6ml 10XTBE Make up to 100ml in distilled water Degas and store at 4°C
$MgCl_2$ (1M)	203.31g $MgCl_2$ Make up to 1l in distilled water and autoclave
Nuseive gel (2%)	2g Nuseive in 100ml 1XTAE buffer and melt in a microwave oven Add 3µl 1% ethidium bromide and mix Leave to cool and pour into a horizontal plate and allow to set
Neutralising solution for dot blots	1.7ml acetic acid in 100ml
Polyacrylamide gel for sequencing	4.6% Acrylamide-bisacrylamide (10.6ml of 40% stock) 1xTBE (filtered) (10ml of 10xTBE stock) 36g Urea Make up to 100ml with water (millipore) Degas, filter and store at 4°C
Post-hybridisation wash solutions for dot blots	
a) 2X wash solution	0.1% SDS (1ml of 10% stock) Make up to 100ml with 2x SSC
b) 0.5X wash solution	0.1% SDS (1ml of 10% stock) 0.5xSSC (25ml of 2x SSC) Make up to 100ml with distilled water
c) 0.1X wash solution	0.1% SDS (1ml of 10% stock) 0.1x SSC (500µl 20x SSC) Make up to 100ml with distilled water

Proteinase-K mix	1% SDS (0.01ml 10% SDS) 2mM EDTA (0.004ml of 0.5M stock) 2mg Proteinase-K Make up to 1ml with distilled water
Saturated sodium chloride	40g NaCl in 100ml distilled water until saturated. Before use agitate and let the NaCl precipitate out
10% SDS	100g SDS dissolved in 1l autoclaved water (work with SDS in the fume hood)
Sodium acetate (3M)	408.24g sodium acetate in distilled water pH 4.6 with acetic acid Make up to 1l
20XSSC	0.15M NaCl (175.35g) 0.15M Na citrate (88.25g) Make up to 1l in distilled water and autoclave
Sucrose-Triton-X lysing buffer	10mM Tris-HCl, pH8 (10ml of 1M stock) 5mM $MgCl_2$ (5ml of 1M stock) 10ml Triton-X 100 Make up to 1l with distilled water Autoclave and add 109.5g sucrose just before use and chill at 4°C
50XTAE	242g Tris base 57.1ml Glacial acetic acid 100ml 0.5M EDTA (pH8) Make up to 1l in distilled water and autoclave Dilute to get a 1x solution before use
10XTBE	108g Tris Base 55g boric acid 9.3g Na_2EDTA Dissolve in distilled water and make up to 1l Autoclave. Dilute ten fold with distilled water before use
T20E5	20mM Tris-HCl (20ml of 1M stock) 5mM EDTA (10ml of 0.5M stock) pH 8 Make up to 1l with water and autoclave

1XTE	10mM Tris (10ml of 1mM stock) 1mM EDTA (2ml of 0.5M stock) pH 8 with HCl Make up to 1l in distilled water
Tris buffer (1M)	121.1g Tris base in distilled water pH 8 with HCl Make up to 1l

APPENDIX 3: Source of reagents and kits

Reagent	Supplier
Acetic acid	Associated Chemical Enterprises
40% Acrylamide-Bisacrilamide	Laboratory specialist services
Ammonium persulphate	Stratagene
Boric acid	Merck
Bromophenol blue sodium salt	Merck
Cronex ® 10s, medical X-ray film	Du Pont
EDTA	Fluka
Ethanol	Merck (BDH)
Ethidium bromide	Sigma
Formamide	Fluka
HGT agarose	Whitehead scientific
MDE	FMC Bioproducts
$MgCl_2$	Merck
NaAc	BDH Chemicals
NaCl	Sky Chemicals
NaOH	SMM Chemicals
Nuseive agarose	Thomson research supplies
Proteinase K	Boehringer Mannheim
SDS	BDH Chemicals
Sephadex G-50	Fluka (Sigma)
Silicon gel slick	FMC Bioproducts
Spermedine	Sigma
Sucrose	Amersham
Temed	Whitehead Scientific
Tris base	SMM Chemicals
Tris-HCl	UnivAR
Triton X-100	Sigma
Urea	Laboratoria
Whatmann 3Chr paper	Merck

Positively charged nylon membrane	Boehringer Mannheim
DIG detection kit	Boehringer Mannheim
CDP-Star	Boehringer Mannheim
Shrimp alkaline phophatase and exonuclease 1	Amersham
Sequencing kits with AmpliTaq DNA polymerase, FS: • dRhodamine dye terminator cycle sequencing ready reaction kit • Big dye terminator cycle sequencing ready reaction kit	Perkin Elmer Applied Biosystems
Taq polymerase and PCR buffer	Boeringer Mannheim
Restriction endonucleases and buffers	Boehringer Mannheim and Ammersham
Wizard miniprep purification kit	Promega
QIAquick gel purification kit	Quiagen

APPENDIX 4: *MC1R, TYR* and *TYRP1* gene consensus sequences

Below are the coding consensus sequences in the 5' to 3' direction of the *MC1R*, *TYR* and *TYRP1* genes used in this project. The positions at which variants were observed have been highlighted. The codon number and nucleotide number for the first codon and nucleotide for each line, respectively, are on the left side of the page.

MC1R gene consensus sequence of the coding region:

The human *MC1R* sequence (GenBank accession number X65634) [Mountjoy *et al* (1992)].

1	M	A	V	Q	G	S	Q	R	R	L	L	G	S	L	N	S	
1	ATG	GCT	GTG	CAG	GGA	TCC	CAG	AGA	AGA	CTT	CTG	GGC	TCC	CTC	AAC	TCC	ACC
18	P	T	A	I	P	Q	L	G	L	A	A	N	Q	T	G	A	R
52	CCC	ACA	GCC	ATC	CCC	CAG	CTG	GGG	CTG	GCT	GCC	AAC	CAG	ACA	GGA	GCC	CGG
35	C	L	E	V	S	I	S	D	G	L	F	L	S	L	G	L	V
103	TGC	CTG	GAG	GTG	TCC	ATC	TCT	GAC	GGG	CTC	TTC	CTC	AGC	CTG	GGG	CTG	GTG
52	S	L	V	E	N	A	L	V	V	A	T	I	A	K	N	R	N
154	AGC	TTG	GTG	GAG	AAC	GCG	CTG	GTG	GTG	GCC	ACC	ATC	GCC	AAG	AAC	CGG	AAC
69	L	H	S	P	M	Y	C	F	I	C	C	L	A	L	S	D	L
205	CTG	CAC	TCA	CCC	ATG	TAC	TGC	TTC	ATC	TGC	TGC	CTG	GCC	TTG	TCG	GAC	CTG
86	L	V	S	G	S#	N	V	L	E	T	A	V	I	L	L	L	E
256	CTG	GTG	AGC	GGG	AGC	AAC	GTG	CTG	GAG	ACG	GCC	GTC	ATC	CTC	CTG	CTG	GAG
103	A	G	A	L	V	A	R	A	A	V	L	Q	Q	L	D	N	V
307	GCC	GGT	GCA	CTG	GTG	GCC	CGG	GCT	GCG	GTG	CTG	CAG	CAG	CTG	GAC	AAT	GTC
120	I	D	V	I	T	C	S	S	M	L	S	S	L	C	F	L	G
358	ATT	GAC	GTG	ATC	ACC	TGC	AGC	TCC	ATG	CTG	TCC	AGC	CTC	TGC	TTC	CTG	GGC
137	A	I	A	V	D	R	Y	I	S	I	F	Y	A	L	R	Y	H
409	GCC	ATC	GCC	GTG	GAC	CGC	TAC	ATC	TCC	ATC	TTC	TAC	GCA	CTG	CGC	TAC	CAC

154	S	I	V	T	L	P	R	A	R*	R[&]	A	V	A	A	I	W	V
460	AGC	ATC	GTG	ACC	CTG	CCG	CGG	GCG	CGG	CGA	GCC	GTT	GCG	GCC	ATC	TGG	GTG
171	A	S	V	V	F	S	T	L	F	I	A	Y	Y	D	H	V	A
511	GCC	AGT	GTC	GTC	TTC	AGC	ACG	CTC	TTC	ATC	GCC	TAC	TAC	GAC	CAC	GTG	GCC
188	V	L	L	C	L	V	V	F	F	L	A	M	L	V	L	M	A
562	GTC	CTG	CTG	TGC	CTC	GTG	GTC	TTC	TTC	CTG	GCT	ATG	CTG	GTG	CTC	ATG	GCC
205	V	L	Y	V	H	M	L	A	R	A	C	Q	H	A	Q	G	I
613	GTG	CTG	TAC	GTC	CAC	ATG	CTG	GCC	CGG	GCC	TGC	CAG	CAC	GCC	CAG	GGC	ATC
222	A	R	L	H	K	R	Q	R	P	V	H	Q	G	F	G	L	K
664	GCC	CGG	CTC	CAC	AAG	AGG	CAG	CGC	CCG	GTC	CAC	CAG	GGC	TTT	GGC	CTT	AAA
239	G	A	V	T	L	T	I	L	L	G	I	F	F	L	C	W	G
715	GGC	GCT	GTC	ACC	CTC	ACC	ATC	CTG	CTG	GGC	ATT	TTC	TTC	CTC	TGC	TGG	GGC
256	P	F	F	L	H	L	T	L	I	V	L	C	P	E	H	P	T
766	CCC	TTC	TTC	CTG	CAT	CTC	ACA	CTC	ATC	GTC	CTC	TGC	CCC	GAG	CAC	CCC	ACG
273	C	G	C	I	F	K	N	F	N	L	F	L	A	L	I	I	C
817	TGC	GGC	TGC	ATC	TTC	AAG	AAC	TTC	AAC	CTC	TTT	CTC	GCC	CTC	[illegible]	ATC	TGC
290	N	A	I	I	D	P	L	I	Y	A	F	H	S	Q	E	L	R
868	AAT	GCC	ATC	ATC	GAC	CCC	CTC	ATC	TAC	GCC	TTC	CAC	AGC	CAG	GAG	CTC	CGC
307	R	T	L	K	E	V	L	T	C	S	W						
916	AGG	ACG	CTC	AAG	GAG	GTG	CTG	ACA	TGC	TCC	TGG						

Changes to consensus sequence: the T90S[#] and P162R* from the sequence in Mountjoy *et al* (1992) and the Q163R[&] change to the sequence from Chhajlani and Wikberg (1992).

TYR gene consensus sequence of the coding region:

The human *TYR* sequence (GenBank accession number M60296) [Giebel *et al*, 1991d].

1	M	L	L	A	V	L	Y	C	L	L	W	S	F	Q	T	S	A
1	ATG	CTC	CTG	GCT	GTT	TTG	TAC	TGC	CTG	CTG	TGG	AGT	TTC	CAG	ACC	TCC	GCT
18	G	H	F	P	R	A	C	V	S	S	K	N	L	M	E	K	E
52	GGC	CAT	TTC	CCT	AGA	GCC	TGT	GTC	TCC	TCT	AAG	AAC	CTG	ATG	GAG	AAG	GAA
35	C	C	P	P	W	S	G	D	R	S	P	C	G	Q	L	S	G
103	TGC	TGT	CCA	CCG	TGG	AGC	GGG	GAC	AGG	AGT	CCC	TGT	GGC	CAG	CTT	TCA	GGC
52	R	G	S	C	Q	N	I	L	L	S	N	A	P	L	G	P	Q
154	AGA	GGT	TCC	TGT	CAG	AAT	ATC	CTT	CTG	TCC	AAT	GCA	CCA	CTT	GGG	CCT	CAA
69	F	P	F	T	G	V	D	D	R	E	S	W	P	S	V	F	Y
205	TTT	CCC	TTC	ACA	GGG	GTG	GAT	GAC	CGG	GAG	TCG	TGG	CCT	TCC	GTC	TTT	TAT
86	N	R	T	C	Q	C	S	G	N	F	M	G	F	N	C	G	N
256	AAT	AGG	ACC	TGC	CAG	TGC	TCT	GGC	AAC	TTC	ATG	GGA	TTC	AAC	TGT	GGA	AAC
103	C	K	F	G	F	W	G	P	N	C	T	E	R	R	L	L	V
307	TGC	AAG	TTT	GGC	TTT	TGG	GGA	CCA	AAC	TGC	ACA	GAG	AGA	CGA	CTC	TTG	GTG
120	R	R	N	I	F	D	L	S	A	P	E	K	D	K	F	F	A
358	AGA	AGA	AAC	ATC	TTC	GAT	TTG	AGT	GCC	CCA	GAG	AAG	GAC	AAA	TTT	TTT	GCC
137	Y	L	T	L	A	K	H	T	I	S	S	D	Y	V	I	P	I
409	TAC	CTC	ACT	TTA	GCA	AAG	CAT	ACC	ATC	AGC	TCA	GAC	TAT	GTC	ATC	CCC	ATA
154	G	T	Y	G	Q	M	K	N	G	S	T	P	M	F	N	D	I
613	GGG	ACC	TAT	GGC	CAA	ATG	AAA	AAT	GGA	TCA	ACA	CCC	ATG	TTT	AAC	GAC	ATC
171	N	I	Y	D	L	F	V	W	M	H	Y	Y	V	S	M	D	A
511	AAT	ATT	TAT	GAC	CTC	TTT	GTC	TGG	ATG	CAT	TAT	TAT	GTC	TCA	ATG	GAT	GCA
188	L	L	G	G	S	E	I	W	R	D	I	D	F	A	H	E	A
562	CTG	CTT	GGG	GGA	TCT	GAA	ATC	TGG	AGA	GAC	ATT	GAT	TTT	GCC	CAT	GAA	GCA

205	P	A	F	L	P	W	H	R	L	F	L	L	R	W	E	Q	E
613	CCA	GCT	TTT	CTG	CCT	TGG	CAT	AGA	CTC	TTC	TTG	TTG	CGG	TGG	GAA	CAA	GAA
222	I	Q	K	L	T	G	D	E	N	F	T	I	P	Y	W	D	W
664	ATC	CAG	AAG	CTG	ACA	GGA	GAT	GAA	AAC	TTC	ACT	ATT	CCA	TAT	TGG	GAC	TGG
239	R	D	A	E	K	C	D	I	C	T	D	E	Y	M	G	G	Q
715	CGG	GAT	GCA	GAA	AAG	TGT	GAC	ATT	TGC	ACA	GAT	GAG	TAC	ATG	GGA	GGT	CAG
256	H	P	T	N	P	N	L	L	S	P	A	S	F	F	S	S	W
766	CAC	CCC	ACA	AAT	CCT	AAC	TTA	CTC	AGC	CCA	GCA	TCA	TTC	TTC	TCC	TCT	TGG
273	Q	I	V	C	S	R	L	E	E	Y	N	S	H	Q	S	L	C
817	CAG	ATT	GTC	TGT	AGC	CGA	TTG	GAG	GAG	TAC	AAC	AGC	CAT	CAG	TCT	TTA	TGC
290	N	G	T	P	E	G	P	L	R	R	N	P	G	N	H	D	K
868	AAT	GGA	ACG	CCC	GAG	GGA	CCT	TTA	CGG	CGT	AAT	CCT	GGA	AAC	CAT	GAC	AAA
307	S	R	T	P	R	L	P	S	S	A	D	V	E	F	C	L	S
919	TCC	AGA	ACC	CCA	AGG	CTC	CCC	TCT	TCA	GCT	GAT	GTA	GAA	TTT	TGC	CTG	AGT
324	L	T	Q	Y	E	S	G	S	M	D	K	A	A	N	F	S	F
970	TTG	ACC	CAA	TAT	GAA	TCT	GGT	TCC	ATG	GAT	AAA	GCT	GCC	AAT	TTC	AGC	TTT
341	R	N	T	L	E	G	F	A	S	P	L	T	G	I	A	D	A
1020	AGA	AAT	ACA	CTG	GAA	GGA	TTT	GCT	AGT	CCA	CTT	ACT	GGG	ATA	GCG	GAT	GCC
358	S	Q	S	S	M	H	N	A	L	H	I	Y	M	N	G	T	M
1072	TCT	CAA	AGC	AGC	ATG	CAC	AAT	GCC	TTG	CAC	ATC	TAT	ATG	AAT	GGA	ACA	ATG
375	S	Q	V	Q	G	S	A	N	D	P	I	F	L	L	H	H	A
1123	TCC	CAG	GTA	CAG	GGA	TCT	GCC	AAC	GAT	CCT	ATC	TTC	CTT	CTT	CAC	CAT	GCA
392	F	V	D	S	I	F	E	Q	W	L	R	R	H	R	P	L	Q
1174	TTT	GTT	GAC	AGT	ATT	TTT	GAG	CAG	TGG	CTC	CGA	AGG	CAC	CGT	CCT	CTT	CAA
409	E	V	Y	P	E	A	N	A	P	I	G	H	N	R	E	S	Y
1125	GAA	GTT	TAT	CCA	GAA	GCC	AAT	GCA	CCC	ATT	GGA	CAT	AAC	CGG	GAA	TCC	TAC

426	M	V	P	F	I	P	L	Y	R	N	G	D	F	F	I	S	S
1276	ATG	GTT	CCT	TTT	ATA	CCA	CTG	TAC	AGA	AAT	GGT	GAT	TTC	TTT	ATT	TCA	TCC
443	K	D	L	G	Y	D	Y	S	Y	L	Q	D	S	D	P	D	S
1327	AAA	GAT	CTG	GGC	TAT	GAC	TAT	AGC	TAT	CTA	CAA	GAT	TCA	GAC	CCA	GAC	TCT
460	F	Q	D	Y	I	K	S	Y	L	E	Q	A	S	R	I	W	S
1378	TTT	CAA	GAC	TAC	ATT	AAG	TCC	TAT	TTG	GAA	CAA	GCG	AGT	CGG	ATC	TGG	TCA
477	W	L	L	G	A	A	M	V	G	A	V	L	T	A	L	L	A
1429	TGG	CTC	CTT	GGG	GCG	GCG	ATG	GTA	GGG	GCC	GTC	CTC	ACT	GCC	CTG	CTG	GCA
494	G	L	V	S	L	L	C	R	H	K	R	K	Q	L	P	E	E
1480	AGA	AAG	CAG	CTT	CCT	GAA	GGG	CTT	GTG	AGC	TTG	CTG	TGT	CGT	CAC	AAG	GAA
511	K	Q	P	L	L	M	E	K	E	D	Y	H	S	L	Y	Q	S
1531	AAG	CAG	CCA	CTC	CTC	ATG	GAG	AAA	GAG	GAT	TAC	CAC	AGC	TTG	TAT	CAG	AGC
528	H	L															
1582	CAT	TTA															

***TYRP1* gene consensus sequence of the coding region:**

The human *TYRP1* sequence (GenBank accession number AF001295) [Box *et al*, 1998].

1	M	S	A	P	K	L	L	S	L	G	C	I	F	F	P	L	L
1	ATG	AGT	GCT	CCT	AAA	CTC	CTC	TCT	CTG	GGC	TGT	ATC	TTC	TTC	CCC	TTG	CTA
18	L	F	Q	Q	A	R	A	Q	F	P	R	Q	C	A	T	V	E
52	CTT	TTT	CAG	CAG	GCC	CGG	GCT	CAA	TTC	CCA	AGA	CAG	TGT	GCC	ACT	GTT	GAG
35	A	L	R	S	G	M	C	C	P	D	L	S	P	V	S	G	P
103	GCT	TTG	AGA	AGT	GGT	ATG	TGT	TGC	CCA	GAC	CTG	TCC	CCT	GTG	TCT	GGG	CCT
52	G	T	D	R	C	G	S	S	S	G	R	G	R	C	E	A	V
154	GGG	ACA	GAC	CGC	TGT	GGC	TCA	TCA	TCA	GGG	AGG	GGC	AGA	TGT	GAG	GCA	GTG
69	T	A	D	S	R	P	H	S	P	Q	Y	P	H	D	G	R	D
205	ACT	GCA	GAC	TCC	CGG	CCC	CAC	AGC	CCT	CAG	TAT	CCC	CAT	GAT	GGC	AGA	GAT
86	D	R	E	V	W	P	L	R	F	F	N	R	T	C	H	C	N
256	GAT	CGG	GAG	GTC	TGG	CCC	TTG	CGC	TTC	TTC	AAT	AGG	ACA	TGT	CAC	TGC	AAC
103	G	N	F	S	G	H	N	C	G	T	C	R	P	G	W	R	G
307	GGC	AAT	TTC	TCA	GGA	CAC	AAC	TGT	GGG	ACG	TGC	CGT	CCT	GGC	TGG	AGA	GGA
120	A	A	C	D	Q	R	V	L	I	V	R	R	N	L	L	D	L
356	GCT	GCC	TGT	GAC	CAG	AGG	GTT	CTC	ATA	GTC	AGG	AGA	AAT	CTT	CTG	GAC	TTA
137	S	K	E	E	K	N	H	F	V	R	A	L	D	M	A	K	R
409	AGT	AAA	GAA	GAA	AAG	AAC	CAC	TTT	GTC	CGG	GCC	CTG	GAT	ATG	GCA	AAG	CGC
154	T	T	H	P	L	F	V	I	A	T	R	R	S	E	E	I	L
460	ACA	ACT	CAC	CCT	TTA	TTT	GTC	ATT	GCC	ACC	AGG	AGA	TCA	GAA	GAA	ATA	CTG
171	G	P	D	G	N	T	P	Q	F	E	N	I	S	I	Y	N	Y
511	GGG	CCA	GAT	GGC	AAC	ACG	CCA	CAA	TTT	GAG	AAC	ATT	TCC	ATT	TAT	AAC	TAC
188	F	V	W	T	H	Y	Y	S	V	K	K	T	F	L	G	V	G
562	TTT	GTT	TGG	ACA	CAC	TAT	TAC	TCA	GTC	AAA	AAG	ACT	TTC	CTT	GGG	GTA	GGA

205	Q	E	S	F	G	E	V	D	F	S	H	E	G	P	A	F	L
613	CAG	GAA	AGC	TTT	GGT	GAA	GTG	GAT	TTC	TCT	CAT	GAG	GGA	CCA	GCT	TTT	CTC
222	T	W	H	R	Y	H	L	L	R	L	E	K	D	M	Q	E	M
664	ACA	TGG	CAC	AGG	TAC	CAC	CTC	CTG	CGT	CTG	GAG	AAA	GAC	ATG	CAG	GAA	ATG
2[illegible]	L	Q	E	P	S	F	S	L	P	Y	W	N	F	A	T	G	K
	TTG	CAA	GAG	CCT	TCT	TTC	TCC	CTT	CCT	TAC	TGG	AAT	TTT	GCA	ACG	GGG	AAA
256	N	V	C	D	I	C	T	D	D	L	M	G	S	R	S	N	F
763	AAT	GTC	TGT	GAT	ATC	TGC	ACG	GAT	GAC	TTG	ATG	GGA	TCC	AGA	AGC	AAC	TTT
273	D	S	T	L	I	S	P	N	S	V	F	S	Q	W	R	V	V
817	GAT	TCC	ACT	CTA	ATA	AGC	CCA	AAC	TCT	GTC	TTT	TCT	CAA	TGG	CGA	GTG	GTC
290	C	D	S	L	E	D	Y	D	T	L	G	T	L	C	N	S	T
868	TGT	GAC	TCC	TTG	GAA	GAT	TAT	GAT	ACC	CTG	GGA	ACA	CTT	TGT	AAC	AGC	ACC
307	E	D	G	P	I	R	R	N	P	A	G	N	V	A	R	P	M
919	GAG	GAT	GGG	CCA	ATT	AGG	AGA	AAT	CCA	GCT	GGA	AAT	GTG	GCC	AGA	CCA	ATG
324	V	Q	R	L	P	E	P	Q	D	V	A	Q	C	L	E	V	G
970	GTG	CAA	CGT	CTT	CCT	GAA	CCA	CAG	GAT	GTC	GCT	CAG	TGC	TTG	GAA	GTT	GGT
341	L	F	D	T	P	P	F	Y	S	N	S	T	N	S	F	R	N
1021	TTA	TTT	GAC	ACG	CCT	CCT	TTT	TAT	TCC	AAC	TCT	ACA	AAC	AGT	TTC	CGA	AAC
358	T	V	E	G	Y	S	D	P	T	G	K	Y	D	P	A	V	R
1072	ACA	GTG	GAA	GGT	TAC	AGT	GAC	CCC	ACG	GGA	AAG	TAT	GAC	CCT	GCT	GTT	CGA
375	S	L	H	N	L	A	H	L	F	L	N	G	T	G	G	Q	T
1123	AGT	CTT	CAC	AAT	TTG	GCT	CAT	CTA	TTC	CTG	AAT	GGA	ACA	GGG	GGA	CAA	ACC
392	H	L	S	P	N	D	P	I	F	V	L	L	H	T	F	T	D
1174	CAT	TTG	TCT	CCA	AAT	GAT	CCT	ATT	TTT	GTC	CTC	CTG	CAC	ACC	TTC	ACA	GAT
409	A	V	F	D	E	W	L	R	R	Y	N	A	D	I	S	T	F
1225	GCA	GTC	TTT	GAT	GAA	TGG	CTG	AGG	AGA	TAC	AAT	GCT	GAT	ATA	TCC	ACA	TTT

426	P	L	E	N	A	P	I	G	H	N	R	Q	Y	N	M	V	P
1276	CCA	TTG	GAA	AAT	GCC	CCT	ATT	GGA	CAT	AAT	AGA	CAA	TAC	AAC	ATG	GTG	CCA
443	F	W	P	P	V	T	N	T	E	M	F	V	T	A	P	D	N
1327	TTC	TGG	CCC	CCA	GTC	ACC	AAC	ACA	GAA	ATG	TTT	GTT	ACT	GCT	CCA	GAC	AAC
460	L	G	Y	T	Y	E	I	Q	W	P	S	R	E	F	S	V	P
1378	CTG	GGA	TAC	ACT	TAT	GAA	ATT	CAA	TGG	CCA	AGT	CGG	GAG	TTT	AGT	GTA	CCT
477	E	I	I	A	I	A	V	V	G	A	L	L	L	V	A	L	I
1435	GAG	ATA	ATT	GCC	ATA	GCA	GTA	GTT	GGC	GCT	TTG	TTA	CTG	GTT	GCA	CTC	ATT
494	F	G	T	A	S	Y	L	I	R	A	R	R	S	M	D	E	A
1480	TTT	GGG	ACT	GCT	TCT	TAT	CTG	ATT	CGT	GCC	AGA	CGC	AGT	ATG	GAT	GAA	GCT
511	N	Q	P	L	L	T	D	Q	Y	Q	C	Y	A	E	E	Y	E
1531	AAC	CAG	CCT	CTC	CTC	ACT	GAT	CAG	TAT	CAA	TGC	TAT	GCT	GAA	GAA	TAT	GAA
528	K	L	Q	N	P	N	Q	S	V	V							
1582	AAA	CTC	CAG	AAT	CCT	AAT	CAG	TCT	GTG	GTC							

APPENDIX 5: Arlequin input files

The input files below were used for haplotype frequency calculations, calculation of the exact test of population differentiation and Hardy-Weinberg equilibrium calculations. Note that '1' at a position indicates the normal sequence and '2' at a position indicates a variant. Each position is a position that a variant was observed in each locus, e.g. position 47 in the *MC1R* gene. Therefore, each individual was either normal at that position (S47 or 1in the input file) or had the variation at that position (I47 or 2 in the input file).

MC1R locus: Haplotypic data for Negroid and San populations

A total of 20 variants were observed at the MC1R locus in the Negroid, San and Caucasoid red-haired individuals. These were at positions 47, 50, 60, 83, 99, 103, 106, 142, 151, 152, 160, 168, 171, 196, 233, 256, 265, 294, 300 and 314. The twenty positions are analysed in this order in the haplotypes and genotypes (Tables 3.1 to 3.4).

```
[Profile]

Title="MC1RNegroidSanHplotypes"
NbSamples= 2
DataType= STANDARD
GenotypicData= 0
LocusSeparator= WHITESPACE
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'

[Data]

[[Samples]]

      SampleName="Negroid"
      SampleSize= 59
      SampleData= {

Ht1   23    11111111111111111111
Ht2   1     21111111111111111112
Ht3   18    11111111111111111112
Ht4   5     11111121111111111112
Ht5   7     11111111111112111111
Ht6   1     11111111111111211111
Ht7   1     11111111111111112111
Ht8   2     11111111111111111121
Ht9   1     11111211111211111112
}

      SampleName="San"
      SampleSize=32
      SampleData={

Ht1   26    11111111111111111111
```

```
Ht3    2      11111111111111111112
Ht7    1      11111111111111112111
Ht10   2      12111111111111111111
Ht11   1      11121111111111111111
}
[[Structure]]
StructureName="Prem"
NbGroups=2
IndividualLevel=0
Group={
"Negroid"
}
Group={
"San"
}
```

MC1R locus: Genotypic data for Negroid, San and Red-haired Caucasoid groups

The positions 1 to 20 referred to in the input file are those listed above.

```
[Profile]

Title="MC1RSanRedNegchildrend"
     NbSamples= 3
     DataType= STANDARD
     GenotypicData= 1
     LocusSeparator= WHITESPACE
     GameticPhase= 0
     RecessiveData= 0
     RecessiveAllele= null
     MissingData= '?'

 [Data]

[[Samples]]

      SampleName="NegroidChildren"
      SampleSize= 22
      SampleData= {

      Gt1   1      11111111111111111111
                   11111111111111111111
      Gt3   9      11111111111111111112
                   11111111111111111111
      Gt4   1      11111111111111111112
                   11111111111111111112
      Gt5   2      11111121111111111112
                   11111111111111111112
      Gt7   1      11111121111111111112
                   11111111111111111111
      Gt13  4      11111111111112111111
                   11111111111111111111
      Gt14  1      11111111111111112111
                   11111111111111111111
      Gt15  1      21111111111111111112
```

```
                    11111111111111111112
        Gt17  1     11111211111211111112
                    11111111111111111112
        Gt29  1     11111111111112111111
                    11111111111111111112
 }

        SampleName="San"
        SampleSize= 17
        SampleData= {

        Gt1   10    11111111111111111111
                    11111111111111111111
        Gt3   2     11111111111111111112
                    11111111111111111111
        Gt14  1     11111111111111112111
                    11111111111111111111
        Gt18  2     12111111111111111111
                    11111111111111111111
        Gt19  1     11121111111111111111
                    11111111111111111111
        Gt20  1     12111111111111112111
                    11111111111111111111
}


        SampleName="CaucasoidRedHeads"
        SampleSize= 7
        SampleData= {

        Gt22  1     11111111112111111111
                    11111111112111111111
        Gt23  1     11111111211111111111
                    11111111211111111111
        Gt24  1     11111111112111111211
                    11111111111111111111
        Gt25  1     11211111121111111111
                    11111111111111111111
        Gt26  1     11112111211111111111
                    11111111111111111111
        Gt27  1     11111111211121111111
                    11111111111111111111
        Gt28  1     11111111211111111111
                    11111111111111121111
}

[[Structure]]

        StructureName = "MC1RSanRedNegchildren"
        NbGroups = 3
        IndividualLevel = 0
        Group ={
        "NegroidChildren"
        }
        Group ={
        "San"
        }
```

```
Group ={
"CaucasoidRedHeads"
}
```

TYR locus: Genotypic data for Negroid and Caucasoid populations

Three positions were analysed here: the 192, 217 and 402 positions, where mutations were found in this project (Table 3.5 and 3.6).

```
[Profile]

Title="TYR"
NbSamples= 2
DataType= STANDARD
GenotypicData= 1
LocusSeparator=TAB
GameticPhase= 0
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'

[Data]

[[Samples]]

SampleName="Negroid"
SampleSize= 14
SampleData= {

TYR1  14    111
            111
}


SampleName="Caucasoid"
SampleSize= 14
SampleData= {

TYR1  5     111
            111
TYR2  5     212
            111
TYR3  1     112
            111
TYR4  1     121
            111
TYR5  2     211
            111
}

[[Structure]]
StructureName="TYR"
NbGroups=2
IndividualLevel=0
```

```
Group={
"Negroid"
}
Group={
"Caucasoid"
}
```

TYRP1 locus: Genotypic data for Negroid and Caucasoid populations

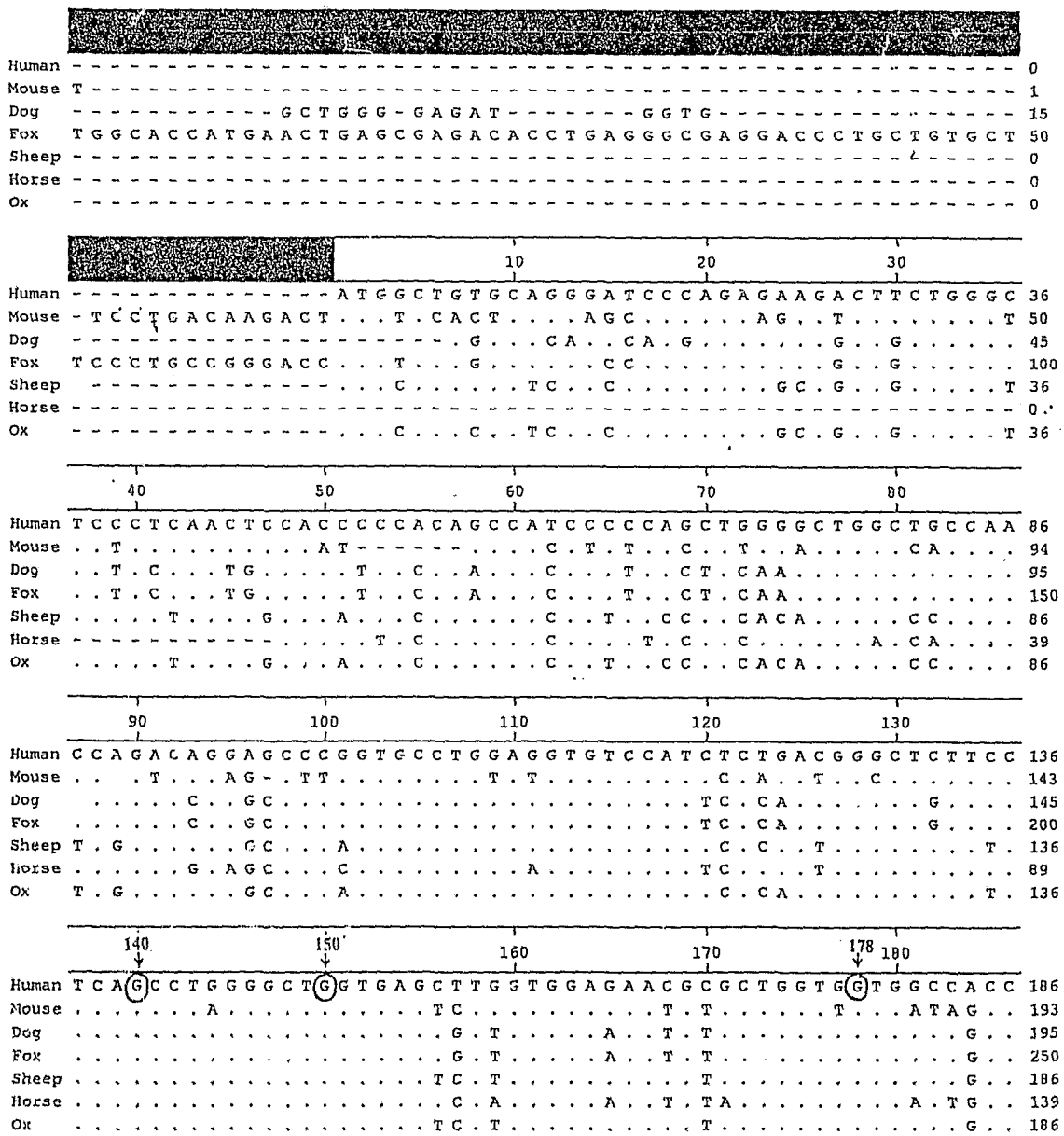
The six positions referred to here are those at which mutations were found in this project. These positions are 7, 87, IVS4-4, 326, IVS7+43 and 519 (Table 3.7 and 3.8).

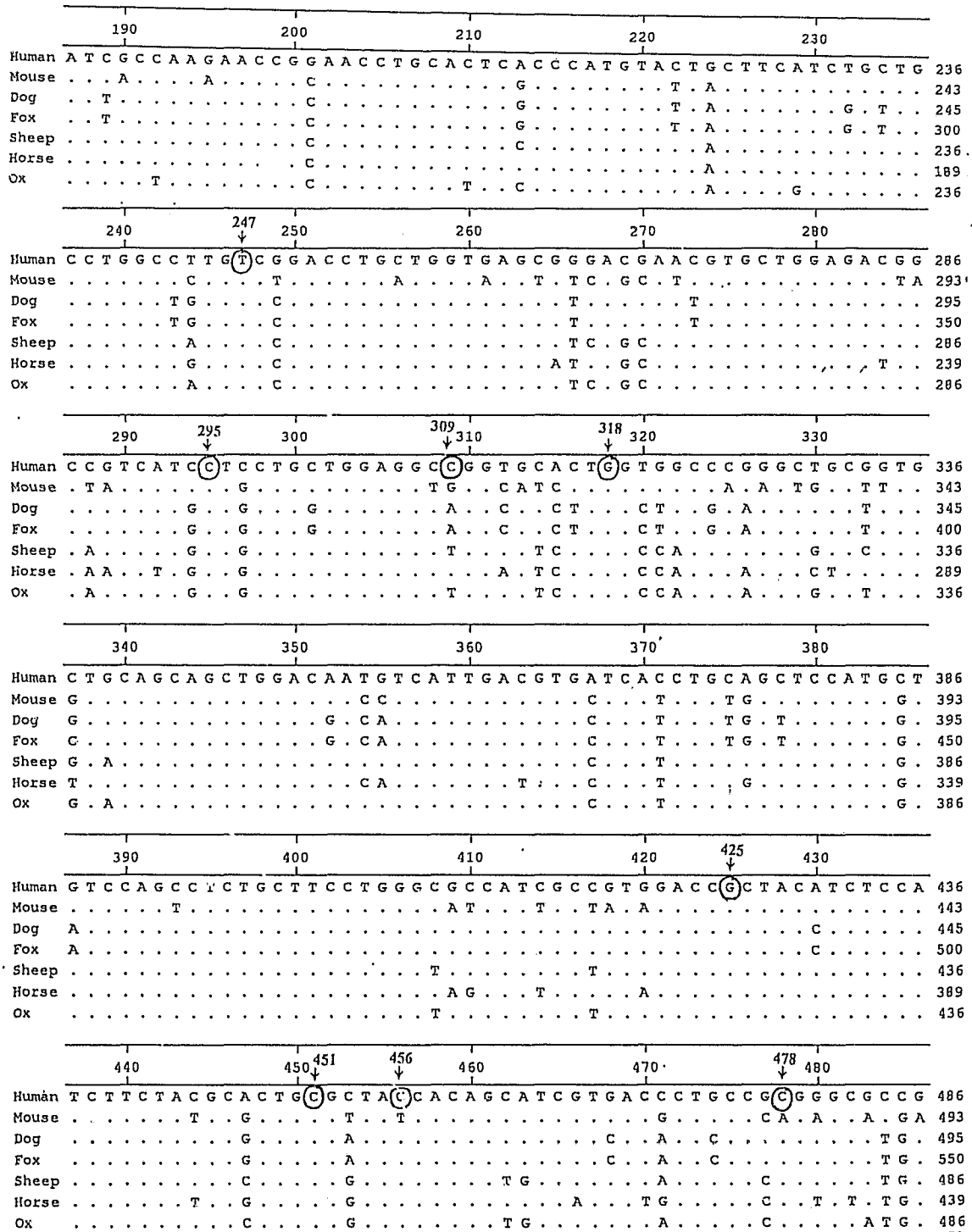
```
[Profile]

Title="TYRP1"
NbSamples= 2
DataType= STANDARD
GenotypicData= 1
LocusSeparator= WHITESPACE
GameticPhase= 1
RecessiveData= 0
RecessiveAllele= null
MissingData= '?'

[Data]

[[Samples]]

SampleName="Negroid"
SampleSize= 15
SampleData= {

TYRP1 7     111111
            111111
TYRP2 1     211111
            111111
TYRP4 1     111211
            111111
TYRP6 4     111121
            111111
TYRP7 2     111112
            111111
}


SampleName="Caucasoid"
SampleSize= 15
SampleData= {

TYRP1 9     111111
            111111
TYRP3 3     121111
            111111
TYRP5 1     112111
            111111
```

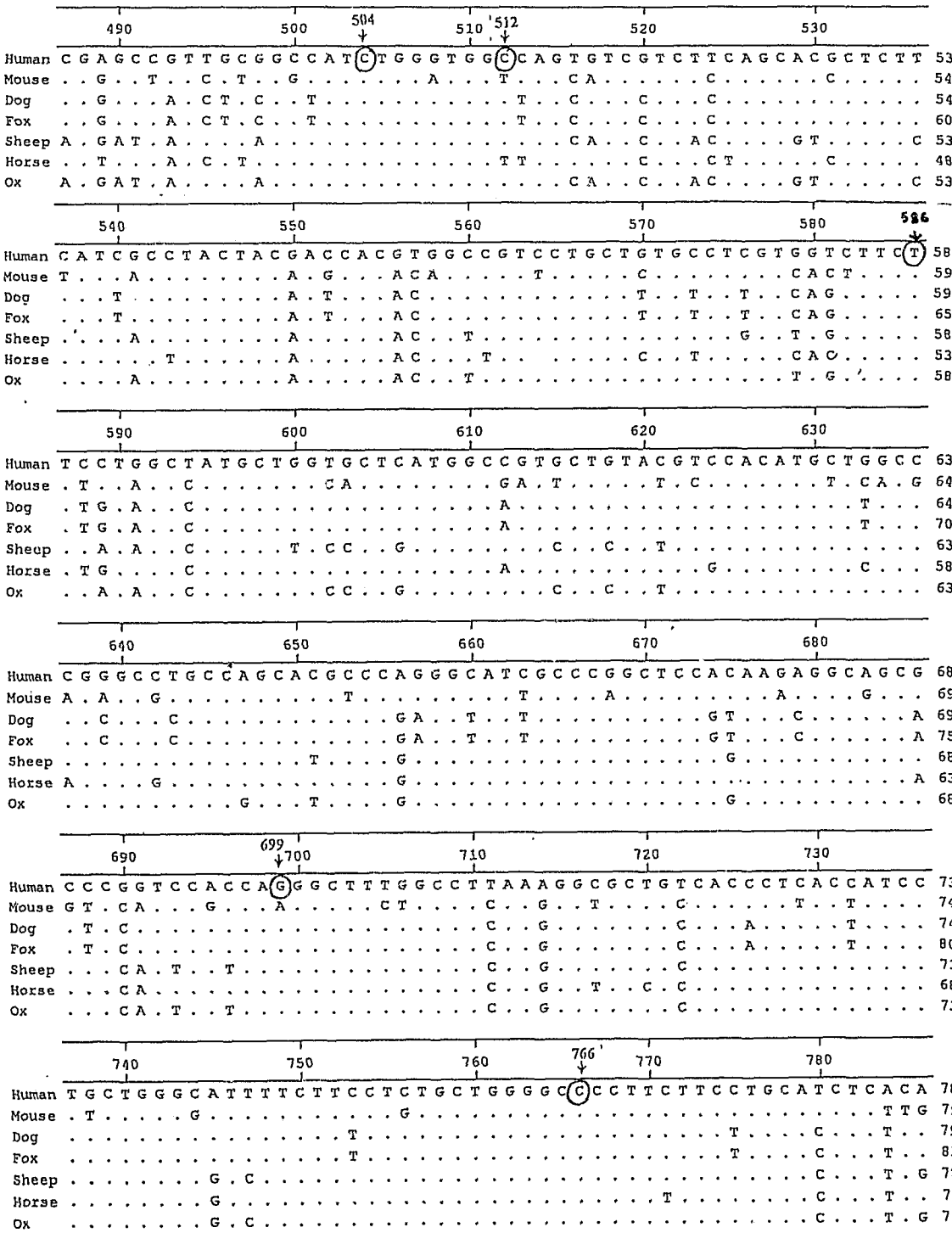
```
TYRP6 2     111121
            111111
}


[[Structure]]
StructureName="TYRP1"
NbGroups=2
IndividualLevel=0
Group={
"Negroid"
}
Group={
"Caucasoid"
}
```

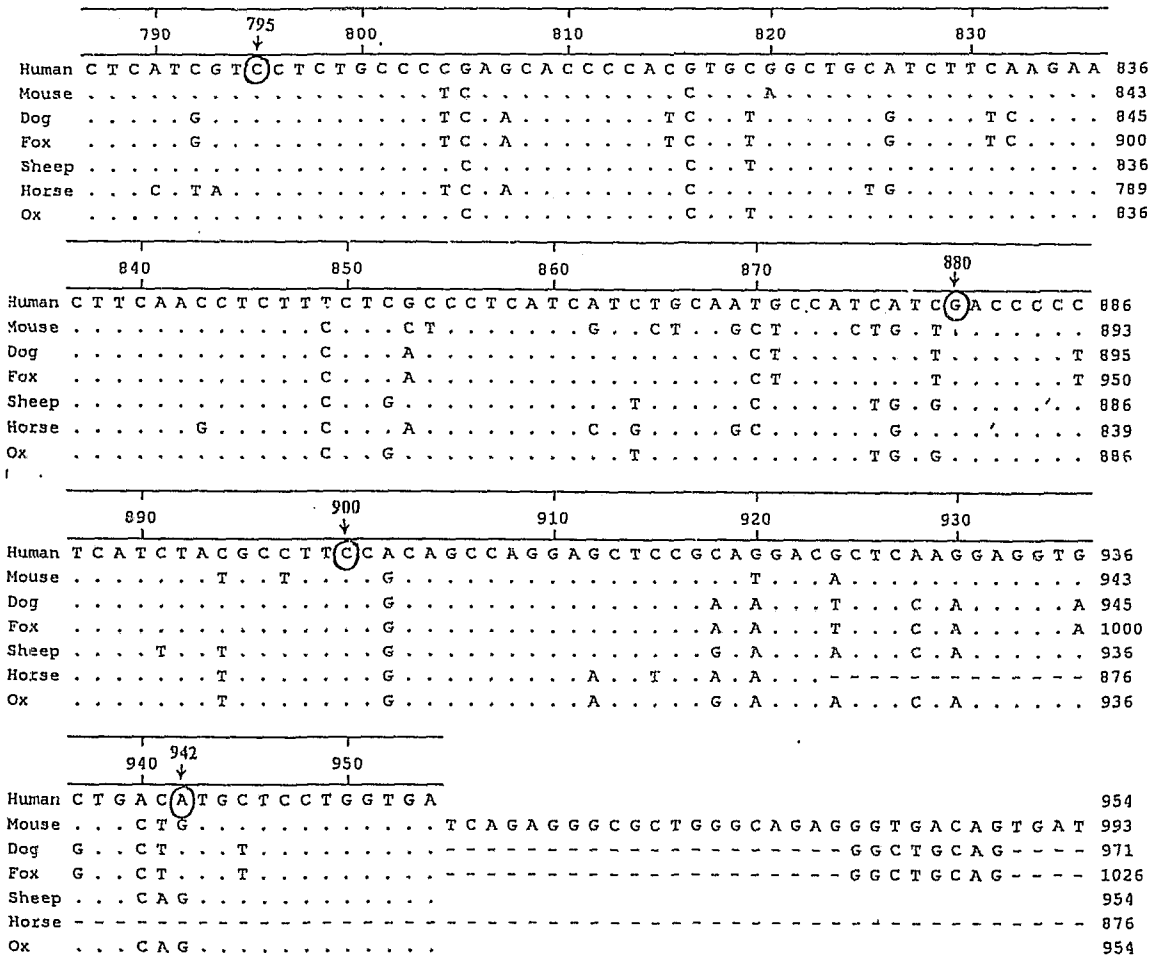
APPENDIX 6: Alignment results for *MC1R*

(a) The human *MC1R* DNA sequence (accession number X65634) was aligned, using the DNASTAR program's MEGALIGN feature, to *MC1R* sequences from: *M. musculus* (mouse) (accession number X65635), *C. familiaris* (dog) (accession number AF064455), *V. vulpes* (red fox) (accession number X90844), *E. caballus* (horse) (accession number X98012), *O. aries* (sheep) (accession number Y13965), and *O. moschatus* (ox) (accession number Y13956). All sequences were obtained from GenBank. The positions at which variants were found in this study have been indicated.




190 200 210 220 230
Human ATCGCCAAGAACCGGAACCTGCACTCACCCATGTACTGCTTCATCTGCTG 236
Mouse ...A....A.....C..........G.........T.A.............. 243
Dog ..T..........C..........G.........T.A.......G.T.. 245
Fox ..T..........C..........G.........T.A.......G.T.. 300
Sheep ..........C..........C..........A.............. 236
Horse ..........C..........A.............. 189
Ox .....T.........C.........T..C..........A....G..... 236
240 247 250 260 270 280
Human CCTGGCCTTGTCGGACCTGCTGGTGAGCGGGACGAACGTGCTGGAGACGG 286
Mouse .......C....T.....A....A..T.TC.GC.T..........TA 293
Dog ......TG....C..............T.....T.............. 295
Fox ......TG....C..............T.....T.............. 350
Sheep .......A....C..............TC.GC.............. 286
Horse .......G....C..............AT..GC.....T.. 239
Ox .......A....C..............TC.GC.............. 286
290 295 300 309 310 318 320 330
Human CCGTCATCCTCCTGCTGGAGGCCGGTGCACTGGTGGCCCGGGCTGCGGTG 336
Mouse .TA.....G.........TG..CATC.........A.A.TG..TT.. 343
Dog .....G..G...G.....A..C..CT...CT..G.A.....T... 345
Fox .....G..G...G.....A..C..CT...CT..G.A.....T... 400
Sheep .A.....G..G..........T....TC....CCA.....G..C... 336
Horse .AA..T.G..G..........A.TC....CCA...A...CT..... 289
Ox .A.....G..G..........T....TC....CCA...A...G..T... 336
340 350 360 370 380
Human CTGCAGCAGCTGGACAATGTCATTGACGTGATCACCTGCAGCTCCATGCT 386
Mouse G..............CC..........C...T...TG.......G. 393
Dog G..............G.CA..........C...T...TG.T.....G. 395
Fox C..............G.CA..........C...T...TG.T.....G. 450
Sheep G.A..........................C...T.............G. 386
Horse T..............CA.....T...C...T...G.........G. 339
Ox G.A..........................C...T.............G. 386
390 400 410 420 425 430
Human GTCCAGCCTCTGCTTCCTGGGCGCCATCGCCGTGGACCGCTACATCTCCA 436
Mouse .....T..............AT...T..TA.A.............. 443
Dog A.....C..... 445
Fox A.....C..... 500
Sheep .....T.....T..... 436
Horse .....AG...T.....A..... 389
Ox .....T.....T..... 436
440 450 451 456 460 470 478 480
Humán TCTTCTACGCACTGCGCTACCACAGCATCGTGACCCTGCCGCGGGCGCCG 486
Mouse .....T..G.....T..T.....G.....CA..A..A.GA 493
Dog .....G.....A.....C..A..C.....TG. 495
Fox .....G.....A.....C..A..C.....TG. 550
Sheep .....C.....G.....TG.....A.....C.....TG. 486
Horse .....T..G.....G.....A...TG.....C..T.T.TG. 439
Ox .....C.....G.....TG.....A.....C.....ATG. 486

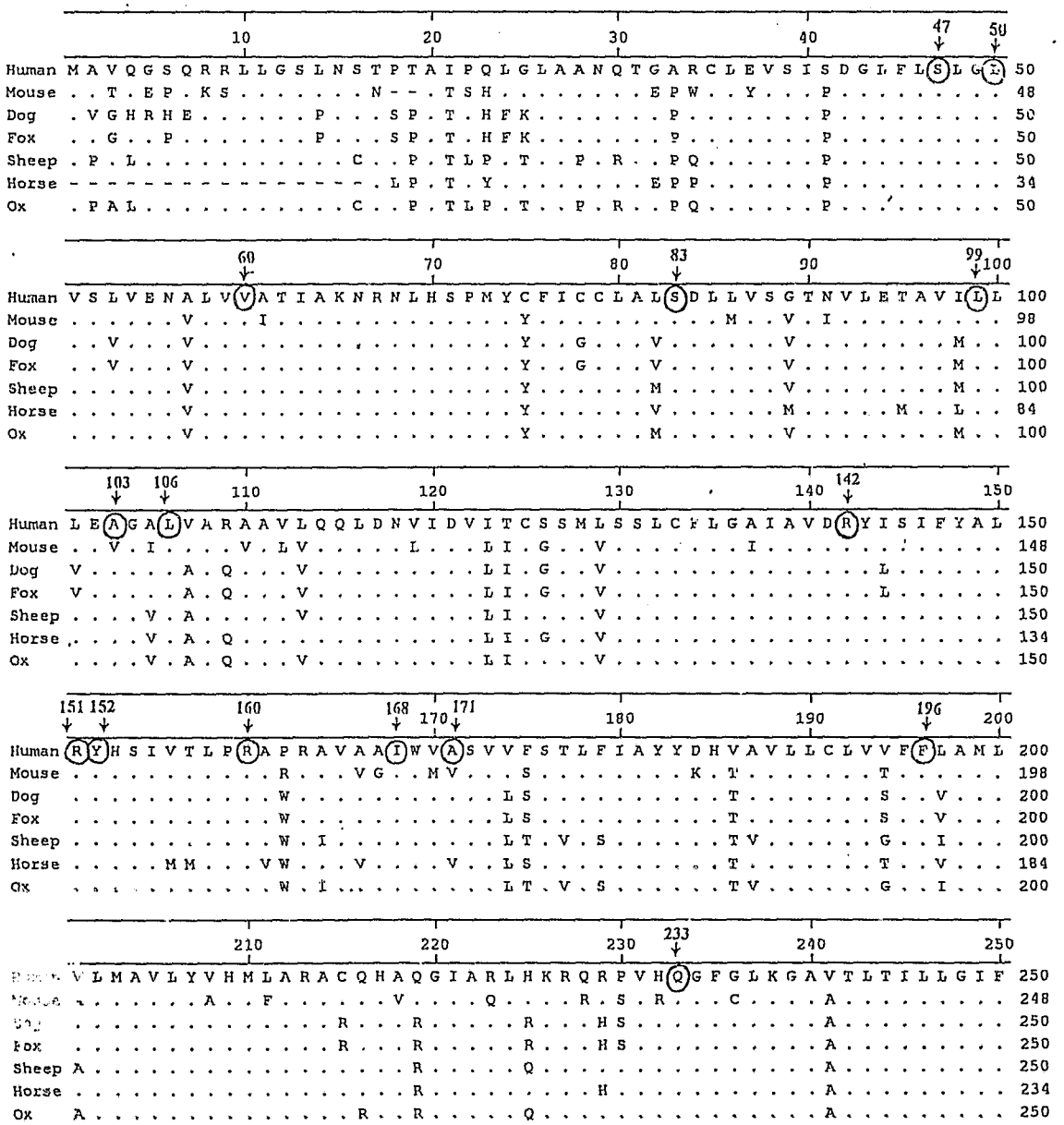

490 500 504 510 512 520 530
Human C G A G C C G T T G C G G C C A T C T G G G T G G C C A G T G T C G T C T T C A G C A C G C T C T T 536
Mouse . . G . . T . . C . T . . G . . . . . . A . . . T . . . C A . . . . . C . . . . . C 543
Dog . . G . . . A . C T . C . . T . . . . . T . . C . . . C . . . C 545
Fox . . G . . . A . C T . C . . T . . . . . T . . C . . . C . . . C 600
Sheep A . G A T . A . . . . A . . . . . C A . . C . . A C . . . . G T . . . . . C 536
Horse . . T . . . A . C . T . . . . . T T . . . . . C . . . C T . . . . . C 489
Ox A . G A T . A . . . . A . . . . . C A . . C . . A C . . . . G T . . . . . C 536
540 550 560 570 580 586
Human C A T C G C C T A C T A C G A C C A C G T G G C C G T C C T G C T G T G C C T C G T G G T C T T C T 586
Mouse T . . . A . . . . A . G . . . A C A . . . . . T . . . . . C . . . . . C A C T 593
Dog . . . T . . . . A . T . . . A C . . . . . T . . T . . T . . C A G 595
Fox . . . T . . . . A . T . . . A C . . . . . T . . T . . T . . C A G 650
Sheep . . . . A . . . . . . A . . . . . A C . . T . . . . . G . . T . G 586
Horse . . . . . T . . . . . A . . . . . A C . . . T . . . . . C . . T . . . . . C A C 539
Ox . . . . A . . . . . A . . . . . A C . . T . . . . . T . G 586
590 600 610 620 630
Human T C C T G G C T A T G C T G G T G C T C A T G G C C G T G C T G T A C G T C C A C A T G C T G G C C 636
Mouse . T . . A . . C . . . . . C A . . . . . G A . T . . . . T . C . . . . . T . C A . G 643
Dog . T G . A . . C . . . . . A . . . . . T . . . 645
Fox . T G . A . . C . . . . . A . . . . . T . . . 700
Sheep . . A . A . . C . . . . . T . C C . . G . . . . . C . . C . . T 636
Horse . T G . . . . C . . . . . A . . . . . G . . . . . C . . . 589
Ox . . A . A . . C . . . . . C C . . G . . . . . C . . C . . T 636
640 650 660 670 680
Human C G G G C C T G C C A G C A C G C C C A G G G C A T C G C C C G G C T C C A C A A G A G G C A G C G 686
Mouse A . A . . G . . . . . T . . . . . T . . . . A . . . . . A . . . . G . . . 693
Dog . . C . . . C . . . . . G A . . T . . T . . . . . G T . . . C . . . . . A 695
Fox . . C . . . C . . . . . G A . . T . . T . . . . . G T . . . C . . . . . A 750
Sheep . . . . . T . . . . G . . . . . G 686
Horse A . . . . G . . . . . G . . . . . A 639
Ox . . . . . G . . . T . . . . G . . . . . G 686
690 699 700 710 720 730
Human C C C G G T C C A C C A G G C T T T G G C C T T A A A G G C G C T G T C A C C C T C A C C A T C C 736
Mouse G T . C A . . . G . . . A . . . . . C T . . . C . . G . . T . . . . C . . . . . T . . T 743
Dog . T . C . . . . . C . . G . . . . . C . . . A . . . . . T 745
Fox . T . C . . . . . C . . G . . . . . C . . . A . . . . . T 800
Sheep . . . C A . T . . T . . . . . C . . G . . . . . C 736
Horse . . . C A . . . . . C . . G . . T . . C . C 689
Ox . . . C A . T . . T . . . . . C . . G . . . . . C 736
740 750 760 766 770 780
Human T G C T G G G C A T T T T C T T C C T C T G C T G G G G C C C C T T C T T C C T G C A T C T C A C A 786
Mouse . T . . . . . G . . . . . G . . . . . T T G 793
Dog . . . . . T . . . . . T . . . . C . . . T . . 795
Fox . . . . . T . . . . . T . . . . C . . . T . . 850
Sheep . . . . . G . C . . . . . C . . . T . G 786
Horse . . . . . G . . . . . T . . . . . C . . . T . . 739
Ox . . . . . G . C . . . . . C . . . T . G 786

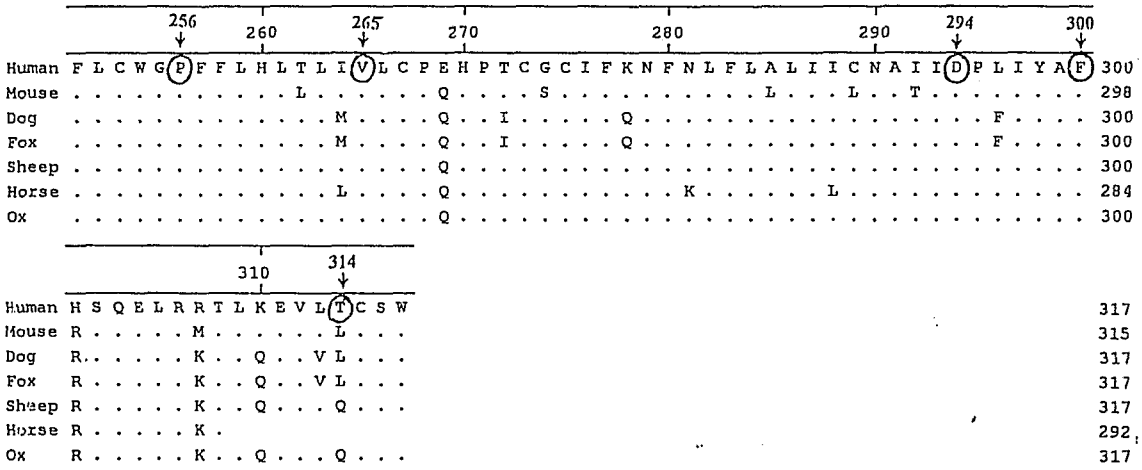


"-" = indicates the absence of residues

"." = indicates the residues that match the human consensus sequence

(b) The human *MC1R* protein sequence (accession number CAA46588) was aligned, using the DNASTAR program's MEGALIGN feature, to *MC1R* sequences from: *M. musculus* (mouse) (accession number CAA46589), *C. familiaris* (dog) (accession number AAC33737), *V. vulpes* (red fox) (accession number CAA62349), *E. caballus* (horse) (accession number CAA66641), *O. aries* (sheep) (accession number CAA74298), and *O. moschatus* (ox) (accession number CAA74290). All sequences were obtained from GenBank. The positions at which variants were found in this study have been indicated.





"-" = indicates the absence of residues

"." = indicates the residues that match the human consensus sequence

REFERENCES

Abdel-Malek Z., Swope V., Smalara D., Babcock G., Dawes S., and Nordland J. (1994). Analysis of the UV-induced melanogenesis and growth of human melanocytes. Pigment Cell Res 7: 326-332.

Abdel-Malek Z., Swope V., Suzuki I., Akcali C., Harriger M. D., Boyce S. T., Urabe K., and Hearing V. J. (1995). Mitogenic and melanogenic stimulation of normal human melanocytes by melanotropic peptides. Proc Natl Acad Sci USA 92: 1789-1793.

Ando H., Itoh A., Mishima Y., and Ichihashi M. (1995). Correlation between the number of melanosomes, tyrosinase mRNA levels, and tyrosinase activity in cultured murine melanoma cells in response to various melanogenesis regulatory agents. J Cellular Phys 163: 608-614.

Avioli L. V. (1979). Disease of the bone: Vitamin D. Part XXII. In: Beeson P. B., McDermott W., and Wyngaarden J. B. (Eds). Cecil text book of medicine. Fifteenth edition. W. B. Saunders Company. Pages 2232-2233.

Awumey E. M., Mitra D. A., Hollis B. W., Kumar R., and Bell N. H. (1998). Vitamin D metabolism is altered in Asian Indians in the southern United States: A clinical research center study. J Clin Endochrinol Metab 83: 169-173.

Bachrach S., Fisher J., and Parks J. S. (1979). An outbreak of vitamin D deficiency rickets in a susceptible population. Paediatrics 64: 871-877.

Baldwin C. T. Hoth C. F., Macina R. A. and Milunsky A. (1995). Mutations in the PAX3 that cause Waardenburg syndrome type I: Ten new mutations and review of the literature. Am J Med Genet 58: 115-122.

Barrat F. J., Auloge L., Pastural E., Lagelouse R. D., Vilmer E., Cant A. J., Weissenbach J. *et al.* (1996). Genetic and physical mapping of the Chediak-Higashi syndrome on chromosome 1q42-43. Am J Hum Genet 59: 625-632.

Barsh G. S. (1996). The genetics of pigmentation: From fancy genes to complex traits. TIG 12: 299-305.

Barton D. E., Kwon B. S., and Francke U. (1988). Human tyrosinase gene, mapped to chromosome 11 (q14→q21), defines second region of homology with mouse chromosome 7. Genomics 3: 17-24.

Bassi M. T., Bergen A. A. B. Bergen., Wapenaar M. C., Schiaffino M. V., Van Schooneveld M., Yates J. R. W., Charles S. J. *et al* (1994). A submicroscopic deletion in a patient with isolated X-linked ocular albinism (OA1). Hum Mol Genet 3: 647-648.

Bennette D. (1991). Colour genes, oncogenes and melanocyte differentiation. J. Cell Sci 98: 135-139.

Boissy R. E. (1988). The melanocyte. Its structure, function, and subpopulations in skin, eyes and hair. Dermatol Cin 6: 161-173.

Boissy R. E., Zhao H., Oetting W. S., Austin L. M., Wildenberg S. C., Boissy Y. L., Zhao Y. *et al* (1996). Mutation in and lack of expression of tyrosinase-related protein-1 (TRP-1) in melanocytes from an individual with brown oculocutaneous albinism: A subtype of albinism classified as "OCA3". Am J Hum Genet 58: 1145-1156.

Boissy R. E., Sakai C., Zhao H., Kobayashi T., and Hearing V. J. (1998). Human tyrosinase related protein-1 (TRP-1) does not function as DHICA oxidase activity in contrast to murine TRP-1. Exp Dermatol 7: 198-204.

Bolognia J. L. and Pawelek J. M. (1988). Continuing medical education. Biology of hypopigmentation. J Am Acad Dermatol 19: 217-255.

Bouchard E., Del Marmol V., Jackson I. J., Cherif D., and Dubertret L. (1994). Molecular characterisation of a human tyrosinase-related-protein-2 cDNA patterns of expression in melanocyte cells. Eur J Biochem 219: 127-134.

Box N. F., Wyeth J. R., O'Gorman L. E., Martin N. G., and Sturm R. A. (1997). Characterisation of melanocyte-stimulating hormone receptor variant alleles in twins with red hair. Hum Mol Genet 6: 1891-1897.

Box N. F., Wyeth J. R., Mayne C. J., O'Gorman L. E., Martin N. G., and Sturm R. A. (1998). Complete sequence and polymorphism study of the human TYRP1 gene encoding tyrosinase-related protein 1. Mamm Genome 9: 50-53.

Bud P. S. and Jackson I. J. (1995). Structure of the mouse tyrosinase-related protein-2/Dopachrome tautomearse (*Tyrp2/Dct*) gene and sequence of two novel slaty genes. Genomics 29: 35-43.

Bultman S. C., Michaud E. J., and Woychik R. P. (1992). Molecular characterisation of the mouse agouti locus. Cell 71: 1195-1204.

Butler M. G. (1989). Hypopigmentation: a common feature of Prader-Willi syndrome. Am J Hum Genet 45: 140-146.

Buxton P. A. (1955). The natural history of the tsetse flies. London School of Hygiene and Tropical Medicine, H K Lewis and Co Ltd.

Cavalli-Sforza L. L., Menozzi P., and Piazza A. (1994). The genetic history of world populations. Chapter 2. In: The history and geography of human genes. Princeton University Press. Pages 60-154.

Chakraborty A. and Pawelek J. (1993). MSH receptor in immortalized human epidermal keratinocytes: A potential mechanism for coordinate regulation of the epidermal-melanin unit. J Cellular Phys 157: 344-350.

Chan C-T. J., Clayton-Smith J., Cheng X-J., Buxton J., Webb T., Pembrey M. E., and Malcolm S. (1993). Molecular mechansims in Angelman syndrome: a survey of 93 patients. J Med Genet 30: 895-902.

Chang A. C. Y., Cochet M., and Cohen S. N. (1980). Structural organisation of human genomic DNA encoding the pro-opiomelanocortin peptide. Proc Natl Acad Sci USA 8: 4890-4894.

Chhajlani V. and Wikberg J. E. S. (1992). Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA. FEBS letters 309: 417-420.

Cochet M., Chang A. C. Y., and Cohen S. N (1982). Characterization of the structural gene and putative 5'-regulatory sequences for human proopiomelanocortin. Nature 297: 335-339.

Cohen T., Muller R. M., Tomita Y., and Shibahara S. (1990). Nucleotide sequence of the cDNA encoding human tyrosinase-related protein. Nucleic Acids Res 18: 2807.

Conner B. J., Reyes A. A., Morin C., Itakura K., Teplitz R. L., and Wallace R. B. (1983). Detection of sickle cell βs-globin gene allele by hybridization with synthetic oligonucleotides. Proc. Natl. Acad Sci. USA 80: 278-282.

Crombie I. K. (1979). Racial differences in melanoma incidence. Br J Cancer 40: 185-193.

Davis M. G. (1991). Electrophoretic techniques. Chapter 7. In: Wilson K. and Goulding K. H. (Eds). Principles and techniques of practical biochemistry. Third Edition. Pages 245-268.

Del Marmol V., Ito S., Jackson I., Vachtenheim J., Berr P., Ghanem G., Morandini R. *et al* (1993). TRP-1 expression correlates with eumelanogenesis in human pigment cells in culture. FEBS letters 327: 307-310.

Durham-Pierre D., King R. A., Naber J. M., Laken S., and Brillian M. H. (1996). Estimation of carrier frequency of a 2.7kb deletion allele of the P-gene associated with OCA2 in African-Americans. Hum Mutation 7: 370-373.

Eberle A. N. (1980). MSH receptors. Chapter 12. In: Schultser D. and Levitzki A. (Eds). Cellular receptors. John Wiley and Sons. Pages 219-231.

Eberle A. N. (1988). The melanotropins: Chemistry, physiology and mechanisms of action. Basal S. Karger publishers.

Eiberg H. and Mohr J. (1987). Major locus for red hair colour linked to MNS blood groups on chromosome 4. Clin Genet 32: 125-128.

Eiberg H. and Mohr J. (1996). Assignment of genes coding for brown eye colour (BEY2) and brown hair colour (HCL3) on chromosome 15q. Eur J Hum Genet 4: 237-241.

Fitzpatrick T. B. (1988). The validity and practicality of sun-reactive skin types I through VI. Arch Dermatol 124: 869-871.

Frandberg P-A, Doufexis M., Kapas S. and Chhaijlani V. (1998). Human pigmentation phenotype: a point mutation generates nonfunctional MSH receptor. Biochem Biophys Res Commun 245: 490-492.

Fraser D. R. (1980). Vitamin D. Chapter 2. In: Barker B. M. and Bender D. A. (Eds.). Vitamins in Medicine. Volume 1. Fourth edition. The Whitefriars Press Limited. Pages 42-146.

Fukai K., Holmes S. A., Lucchese N. J., Siu V. M., Weleber R. G., Schnur R. E. and Spritz R. A. (1995a). Autosomal recessive ocular albinism associated with a functionally significant tyrosinase gene polymorphism. Nature Genet 9: 92-95.

Fukai K. Oh J., Frenk E., Amodovar C., and Spritz R. A. (1995b). Linkage disequilibrium mapping of the gene for Hermansky-Pudlak syndrome to chromosome 10q23.1-q23.3. Hum Mol Genet 4: 1665-1669.

Furumura M., Sakai C., Potterf S. B., Vieira W. D., Garsh G. S., and Hearing V. J. (1998a). Characterization of genes modulated during pheomelanogenesis using differential display. Proc Natl Acad Sci USA 95: 7374-7378.

Furumura M., Solano F., Matsunaga N., Sakai C., Spritz R. A., and Hearing V. J. (1998b). Metal ligand-binding specificities of the tyrosinase-related proteins. Biochem Biophys Res Commun 242: 579-585.

Gahl W. A., Potterf B., Durham-Pierre D., Brilliant M. H., and Hearing V. J. (1995). Melanosomal tyrosine transport in normal and pink-eyed dilution murine melanocytes. Pigment Cell Res 8: 229-233.

Gedde-Dahl T., Oleisen B., Siverts A., and Wilhelmy M (1982). Support for synteny of PTC-K with Jk-IJK-BEY1-Co. Cytogenet Cell Genet 32:278.

Giebel L. B. and Spritz R. A. (1990). RFLP for MboI in the human tyrosinase (TYR) gene detected by PCR. Nucleic Acids Res 18: 3103.

Giebel L. B., Musarella M. A., and Spritz R A. (1991a). A nonsense mutation in the tyrosinase gene of Afghan patients with tyrosinase negative (type IA) oculocutaneous albinism. J Med Genet 28: 464-467.

Giebel L. B., Tripathi R. K., Strunk K. M., Hanifin J. M., Jackson C. E., King R. A., and Spritz R. A. (1991b). Tyrosinase gene mutations associated with type 1B ("yellow") oculocutaneous albinism. Am J Hum Genet 48: 1159-1167.

Giebel L. B., Tripathi R. K., King R. A., and Spritz R. A. (1991c). A tyrosinase gene missense mutation in temp-sensitive type I oculocutaneous albinism- a homologue to the Siamese cat and the Himalayan mouse. J Clin Invest 87: 1119-1122.

Giebel L. B., Strunk K. M., and Spritz R. A. (1991d). Organization and nucleotide sequences of the human tyrosinase gene and a truncated tyrosinase-related segment. Genomics 9: 435-445.

Giebel L. B. and Spritz R. A. (1992). The molecular basis of type 1 (tyrosinase-deficient) human oculocutaneous albinism. Pigment Cell Res (suppl 2): 101-106.

Glenn C. C., Driscoll D. J., Yang T., and Nicholls R. (1997). Genomic imprinting: potential function and mechanisms revealed by the Parder-Willi and Angelman syndromes. Mol Hum Reproduction 3: 321-332.

Grantz I., Yamada T., Tashiro T., Konda Y., Shimoto Y., Miwa H., and Trent J. M. (1994). Mapping of the gene encoding the melanocortin-1 (α-Melanocytes Stimulating Hormone) receptor (MC1R) to human chromosome 16q24.3 by fluorescence *in situ* hybridisation. Genomics 19: 394-395.

Gunay-Aygun M., Cassidy S. B., and Nicholls R. D. (1997). Prader-Willi and other syndromes associated with obesity and mental retardation. Behavior Genet 27: 307-324.

Halaban R. and Moellmann G. (1990). Murine and human *b* locus pigmentation genes encode a glycoprotein (gp75) with catalase activity. Proc Natl Acad Sci USA 87: 4809-4813.

Halaban R. and Moellmann G. (1993). White mutants in mice shedding light on humans. J Invest Dermatol 100: 176S-185S.

Hauge and Helweg-Larsen (1954).Studies of linkage in man. Red hair verses blood groups, P.T.C. and eye colour. Ann Eugen 18: 175-182.

Hayashi K. (1991). PCR-SSCP: a simple sensitive method for detection of mutations in the genomic DNA. PCR methods and applications 1: 34-38.

Hayashi K. and Yandell D. W. (1993). How sensitive is PCR? Hum Mutation 2: 338-346.

Hazelwood S., Shotelersuk V., Wildenberg S. C., Chen D., Iwata F., Kaiser-Kupfer M. I., White J. G. *et al* (1997). Evidence for locus heterogeneity in Puerto Ricans with Hermansky-Pudlak syndrome. Am J Hum Genet 61: 1088-1094.

Hearing V. J. and Jimenez M. (1989). Analysis of mammalian pigmentation at the molecular level. Pigment Cell Res 2: 75-85.

Hearing V. J. and K. Tsukamoto (1991). Enzymatic control of pigmentation in mammals. FASEB J 5: 2902-2909.

Hill H. Z. (1992). The function of melanin or six blind people examine an elephant. Bioessays 14: 49-56.

Hoganson G. E., Ledwitz-Rigby F., Davidson R. L., and Fuller B. B. (1989). Regulation of tyrosinase mRNA levels in mouse melanoma cell clones by melanocyte stimulating hormone cyclic AMP. Somatic Cell Mol Genet 15: 255-263.

Ichii-Jones F., Lear J. T., Heagerty A. H. M., Smith A. G., Hutchinson P. E., Osborne J., Bowers B. *et al.* (1998).Susceptibility to melanoma: influence of skin type and polymorphism in the melanocyte stimulating hormone receptor gene. J Invest Dermatol 111: 218-221.

Innis M. A and Gefald D. H. (1990). Part one. Basic methodology. Chapter 1: Optimisation of PCRs. In: Innis M. A., Gefald D. H. Sninsky J. J. and White T. J. (Eds). PCR protocols: A guide to methods and applications. Pages 36-37.

Jackson I. J., Chambers D. M., Budd P. S., and Johnson R. (1991). The tyrosinase-related protein-1 gene has a structure and promoter sequence very different from tyrosinase. Nucleic Acids Res 19: 3799-3804.

Jackson I. J., Chambers D. M., Tsukamoto K., Copeland N. G., Gilbert D. J., Jenkins N. A., and
Hearing V. J. (1992). A second tyrosinase-related protein, TRP-2, maps to and is mutated at the mouse *slaty* locus. EMBO J 11: 527-535.

Jackson I. J. (1993a). More to colour than meets the eye. Current Biol 8: 518-521.

Jackson I. J. (1993b). Colour coded switches. Nature 362: 587-588.

Jackson I. J., Budd P., Horn J. M., Johnson R., Raymond S., and Steel K. (1994a). Genetics and molecular biology of mouse pigmentation. Pigment Cell Res 7: 75-80.

Jackson I. J. (1994b). Evolution and expression of tyrosinase related proteins. Pigment Cell Res 7: 241-242.

Jimbow K., Ishida O., Ito S., Hori Y., Witkop C. J., and King R. A. (1983). Combined chemical and electron microscopic studies of pheomelanosomes in red hair. J Invest Dermatol 81: 506-511.

Jimenez M., Tsukamoto K., and Hearing V. J. (1991). Tyrosinase from two different loci are expressed by normal and by transformed melanocytes. J Biol Chem 266: 1147-1156.

Jiminez-Cervantes C., Solano F., Kobayashi T., Urabe K., Hearing V. J., Lozano J. A., and Garcia-Borron J. C. (1994). A new enzymatic function in the melanogenic pathway. J Biol Chem 269: 17 993-18 001.

Jones S., Martin R. and Pilbeam D. (1994). The cambridge encyclopaedia of human evolutions. Cambridge University Press.

King R. A. and Witkop C. J. (1976). Hairbulb tyrosinase activity in oculocutaneous albinism. Nature 263: 69-71.

King R. A, Olds D. P., and Witkop C. J. (1978). Characterisation of human tyrosinase: properties normal and albino enzyme. J Invest Dermatol 71: 136-139.

King R. A., Lewis R. A., Townsend D., Zelickson A., Olds D. P. and Brumbaugh J. (1985). Brown oculocutaneous albinism clinical, opthalmological, and biochemical characterisation. Opthalmology 92: 1496-1504.

King R. A., Wirtschafter J. D., Olds D. P., and Brumbaugh J. (1986). Minimal pigment: a new type of oculocutaneous albinism. Clin Genet 29: 42-50.

King R. A. and Oetting W. S. (1992). Molecular basis of type IA (tyrosinase negative) oculocutaneous albinism. Pigment Cell Res 2: 249-253.

King R. A., Wiesner G. L., Townsend D., and White J. G. (1993). Hypopigmentation in Angelman syndrome. Am J Med Genet 46: 40-44.

King R. A., Hearing V. J., Creel D. J., and Oetting W. S. (1995). Albinism. In: Scriver C. R., Beaudet A. L., Sly W. S., and Valle D. (Eds). The metabolic and molecular basis of inherited disorders. New York, McGraw-Hill. Pages 4353-4392.

Kobayashi T., Urabe K., Winder A., Jimenez-Cervantes C., Imokawa G., Brewington T., Solano F. *et al.* (1994a). Tyrosinase related protein 1 (TRP1) functions as a DHICA oxidase in melanin synthesis. EMBO J 13: 5818-5825.

Kobayashi T., Urabe K., Winder A., Tsukamoto K., Brewington T., Imokawa G., Potterf B. *et al.* (1994b). DHICA oxidase activity of TRP1 and interactions with other melanogenic enzymes. Pigment Cell Res 7: 227-234.

Koer A. and Pawelek J. (1982). Mammalian tyrosinase catalyses three reactions in the biosynthesis of melanin. Science 217: 1163-1165.

Koppula S. V., Robbins L. S., Lu D., Baack E., White C. R. Jr, Swanson N. A., and Cone R. D. (1997). Identification of common polymorphisms in the coding sequence of the human MSH receptor (MC1R) with possible biological effects. Hum Mutation 9: 30-36.

Krawczak M., Reiss J., and Cooper D. N. (1992). The mutational spectrum of single base-pair substitutions in mRNA splicing junctions of human genes: cause and consequence. Hum Genet 90: 41-54.

Kromberg J. G. and Jenkins T. (1982). Prevalence of albinism in the South African Negro. S Afr Med J 61: 383-386.

Kromberg J. G. R. (1986). A genetic and psychosocial study of albinism in southern Africa. PhD thesis, University of the Witwatersrand.

Kromberg J. G. R. (1987). Albinism in South Africa. IMFAMA: 3-5, 13.

Kromberg J. G. R., Zwane E., Castle D., and Jenkins T. (1987). Albinism in South African Blacks. The Lancet: 388-389.

Kromberg J. G., Castle D., Zwane E., and Jenkins T. (1989). Albinism and skin cancer in southern Africa. Clin Genet 36: 43-52.

Kromberg J. G. R., Castle D., Zwane E. M., and Jenkins T. (1990). Red or rufous albinism in southern Africa. Ophthal Paediatr Genet 11:229-235.

Krude H., Biebermann H., Luck W., Horn R., Bradant G., and Gruters A. (1998). Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by *POMC* mutations in humans. Nature Genet 19: 155-157.

Lamoreux M. L., Zhou B-K., Rosemblat S., and Orlow S. (1995). The pinkeyed-dilution protein and the eumelanin/pheomelanin switch: in support of a unifying hypothesis. Pigment Cell Res 8: 263-270.

Lee S-T., Nicholls R. D., Phil D., Bundey S., Laxova R., Musarella M., and Spritz R. A. (1994). Mutations of the *P* gene in oculocutaneous albinism, ocular albinism, and Prader-Willi syndrome plus albinism. New Eng J Med 330: 529-534.

Lee S-T., Nicholls R. D., Jong M. T. C., Fukai K., and Spritz R. A. (1995). Organization and sequence of the human P gene and identification of a new family of transport proteins. Genomics 26: 354-363.

Lerner A. B., Fitzpatrick T. B., Calkins E., and Summerson W. H. (1949). Mammalian tyrosinase: Preparation and properties. J Biol Chem 178: 185-195.

Lerner A. B., Fitzpatrick T. B., Calkins E., and Summerson W. H. (1950). Mammalian tyrosinase: The relationship of copper to enzymatic activity. J Biol Chem 187: 793-802.

Li X., Park W-J., Pyeritz R. E., and Jabs E. W. (1995). Effect on splicing of a silent FGFR2 mutation in Crouzon syndrome. Nature Genet 9: 232-233.

Li W-H (1997). Molecular evolution. Sinauer Associates, Inc., Publishers.

Little M. A. and Wolff M. E. (1981). Skin and hair reflectance in women with red hair. Ann Hum Biol 8: 231-241.

Liu W., Qian C., and Francke U. (1997). Silent mutation induces exon skipping of fibrillin-1 gene in Marfan syndrome. Nature Genet 16: 328-329.

Lloyd T., Garry F. L., Manders E. K., and Marks J. G. Jr. (1987). The effect of age and hair colour on human hairbulbs tyrosinase activity. Brit J Dermatol 116: 485-489.

Loomis W. F. (1967). Skin-pigment regulation of vitamin-D biosynthesis in man. Science 157: 501-506.

Lund P. M., Puri N., Durham-Pierre D., King R. A., and Brilliant M. H. (1997). Oculocutaneous albinism in an isolated Tonga community in Zimbabwe. J Med Genet 34: 733-735.

Lyon M. F., King T. R., Gondo Y., Gardner J. M., Nakatsu Y., Eicher E. M., and Brilliant M. H. (1992). Genetic and molecular analysis of recessive alleles at the pink-eyed dilution (*p*) locus of the mouse. Proc Natl Acad Sci USA 89: 6968-6972.

Magenis R. E., Smith L., Nadeau H., Johnson K. R., Mountjoy K. G., and Cone R. (1994). Mapping of the ACTH, MSH, and neural (MC3&MC4) melanocortin receptors in the mouse and human. Mamm Genome 5: 503-508.

Manga P., Kromberg J. G. R , Box N. F., Sturm R. A., Jenkins T., and Ramsay M. (1997). Rufous oculocutaneous albinism in southern African blacks is caused by mutations in the *TYRP1* gene. Am J Hum Genet 61: 1095-1101.

Maxam A. M., and Gilbert W. (1977). A new method of sequencing DNA. Proc Natl Acad Sci 74: 560-564.

Miller S. A., Dykes D. D., and Polesky H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.

Morelli J. G., Yohn J. J., Zekman T., and Norris D. A. (1993). Melanocyte movement *in vitro*: role of matrix proteins and integrin receptors. J Invest Dermatol 101: 605-608.

Morrison R., Mason K., and Frost-Mason S. (1995). A cladistic analysis of the evolutionary relationships of the numbers of the tyrosinase gene family using sequencing data. Pigment Cell Res 7: 388-393, 1995.

Mountjoy K. G., Robbins L. S., Mortrud M. T., and Cone R. D. (1992). The cloning of a family of genes that encode the melanocortin receptors. Science 25: 1248-1254.

Mullis K. and Faloona F. A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalysed chain reaction. Methods in Enzymology 155: 355-360.

Murty V. V. V., Bouchard B., Mathew S., Vijayasaradhi S., and Houghton A. N. (1992). Assignment of the human *TYRP* (*brown*) locus to chromosome region 9p23 by nonradioactive in situ hybridisation. Genomics 13: 227-229.

Neer R. M. (1975). The evolutionary significance of vitamin D, skin pigmentation, and ultraviolet light. Am J Phys Anthropol 43: 409-416.

Newton J. M., Orlow S. L., and Barsh G. S. (1996). Isolation and characterisation of a mouse homolog of the X-linked ocular albinism. Genomics 37: 219-225.

Norris D. A., Horikawa T., and Morelli J. G. (1994). Melanocyte destruction and repopulation in vitiligo. Pigment Cell Res 7: 193-203.

Nurse G. T., Weiner J. S., and Jenkins T. J. (1985). Research monographs on human population biology No.3. The peoples of southern Africa and their affinities. Oxford Science Publications.

Oetting W. S. and King R. A. (1994). Analysis of tyrosinase mutations associated with tyrosinase-related oculocutaneous albinism (OCA1). Pigment Cell Res 7: 285-290.

Oetting W. S., Brilliant M. H., and King R. A. (1996). The clinical spectrum of albinism in humans. Mol Med Today: 330-335.

Oh J., Ho L., Ala-Mello S., Amato D., Armstrong L., Bellucci S., Carakushansky G *et al* (1998). Mutation analysis of patients with Hermansky-Pudlak syndrome: a frameshift hot spot in the HPS gene and apparent locus heterogeneity. Am J Hum Genet 62: 593-598.

Okoro A. N. (1975). Albinism in Nigeria. Br J Dermatol 92: 485-492.

Orita M, Iwahara H, Kanazawa H, Hayashi K, Sekiya T (1989a). Detection of polymorphisms of human DNA by gel electrophoresis as single-stranded conformation polymorphisms. Proc Natl Acad Sci USA 86: 2766-2770.

Orita M, Suzuki Y, Sekiya T and Hayashi K (1989b). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874-879.

Orlow S. J., Boissy R. E., Moran D. J., and Pifko-Hirst S. (1993). Subcellular distribution of tyrosinase and tyrosinase related protein-1: implication for melanosomal biogenesis. J Invest Dermatol 100: 55-64.

Orlow S. J. (1995). Congenital disorders of hypopigmentation. Seminars in Dermatol 14: 27-32.

Ortonne J. P. (1990). The effects of UV exposure on skin melanin pigmentation. J Int Med Res 18: 8C-17C.

Ortonne J. and Prota G. (1993). Hair melanins and hair colour: Ultrastructural and biochemical aspects. J Invest Dermatol 101: 82S-89S.

Oskam L., Hartskeerl R. A., Harmans C. J., De Wit M. Y. L., Jarings G. H., Nicholls R. D., and Klatser P. R. (1995). A 46kDa integral membrane protein from *Mycobacterium leprae* resembles a number of bacterial and mammalian membrane transport protein. Microbiology 141: 1963-1968.

Pavel S. J (1993). Dynamics of melanogenesis intermediates. J Invest Dermatol 100: 162S-165S.

Pawelek J. (1991). After dopachrome? Pigment Cell Res 4: 53-62.

Penrose L. S. (1935). The detection of autosomal linkage in data which consist of pairs of brothers and sisters of unspecified parentage. Ann Eugen 6: 133-138.

Pierard G. E. (1998). EEMCO guidance for the assessment of skin colour. J Eur Acad Dermatol Venereol 10: 1-11.

Prober J. M., Trainor G. L., Dam R. J., Hobbs F. W., Robertson C. W., Zagursky R. J., Cocuzza A. J. *et al* (1987). A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. Science 238: 336-341.

Prota G. (1993). Complex biochemistry of melanogenesis- regulatory mechanisms of melanogenesis beyond the tyrosinase concept. J Invest Dermatol 100: 156S-162S.

Prota G. (1997). Pigment cell research. What directions? Pigment Cell Res 10: 5-11.

Puri N. and Brilliant M. H. (1998). The function of the pink-eyed dilution protein. Pigment Cell Res 11: 174.

Ramsay M. and Jenkins T. (1988). Alpha-globin cluster haplotypes in the Kalahari San and southern African Bantu-speaking Blacks. Am J Hum Genet 43:527-533.

Ramsay M., Colman M., Stevens G., Zwane E., Kromberg J., Farrall M., and Jenkins T. (1992). The tyrosinase-positive oculocutaneous albinism locus maps to chromosome 15q11.2-q12. Am J Hum Genet 51: 879-884.

Ramsay M. (1996). Protein trafficking violations. Nature Genet 14: 242-245.

Rana K. B., Hewett-Emmett D., Jin L., Chang B. H-J., Sambuughin N., Lin M., Watkins S., *et al.* (1999). High polymorphism at the melanocortin 1 receptor locus. Genetics 151: 1547-1557.

Reintgen D. S., McCarty K. M. Jr, Cox E., Seigler H. F. (1982). Malignant melanoma in black American and white American populations. JAMA 248: 1856-1859.

Relethford J. H. (1997). Hemispheric difference in human skin colour. Am J Anthropol 104: 449-457.

Richard I. and Beckmann J. S. (1995). How neutral are synonymous codon mutations? Nature Genet 10: 259.

Rinchik E. M., Bultman S. J., Horsthemke B., Lee S-T., Strunk K. M., Spritz R. A., Avidano K. M. *et al* (1993). A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. Nature 361: 72-76.

Rippey J. J. and Rippey E. (1984). Epidemiology of malignant melanoma of the skin in South Africa. S Afr Med J 65: 595-598.

Robbins L. S., Nadeau J. H., Johnson K. R., Kelly M. A., Roselli-Rehfuss L., Baack E., Mountjoy K. G. *et al.* (1993). Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. Cell 72: 827-834.

Roberts D. F., Kromberg J. G. R., and Jenkins T. (1986). Differentiation of heterozygotes in recessive albinism. J Med Genet 23: 323-327.

Robins A. H. (1991). Biology of the pigment cell. CambridgeUniversity Press.

Rose E. F. (1973). Pigment variation in relation to protection and susceptibility to cancer. Pigment Cell Res 1: 236-245.

Rosenblum B. B., Lee L. G., Spurgeon S. L., Khan S. H., Menchen S. M., Heiner C. R., and Chen S. M. (1997). New dye-labeled terminators for improved DNA sequencing patterns. Nucleic Acids Res. 25: 4500-4504.

Rungta D., Corn T. D., and Fuller B. B. (1996). Regulation of tyrosinase mRNA in mouse melanoma cells by α-melanocyte-stimulating hormone. J Invest Dermatol 107: 689-693.

Saiki R. K., Walsh P. S., Levenson C. H., and Erlich H. A. (1989). Genetic analysis of amplified DNA with immobilized sequence-specific oligonucleotide probes. Proc Natl Acad Sci USA 86: 6230-6234.

Sakai C., Ollmann M., Kobayashi T., Abdel-Malek Z., Muller J., Vieira W., Imokawa G. *et al.* (1997). Modulation of murine melanocyte function *in vitro* by agouti signaling protein. EMBO J 16: 3544-3552.

Sambrook J, Fritsd E F, and Maniatis T (1989). Molecular cloning: A laboratory manual. Book 1. Second edition. Cold Spring Harbour Laboratory Press.

Sanger F., Nicklen S., and Coulson A. R. (1977). DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467.

Shefield V. C., Beck J. S., Kwitek A. E., Sandstrom D. W., and Stone E. M. (1993). The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. Genomics 16: 325-332.

Smith L. M., Sanders J. Z, Kaiser R. J., Hughes P., Dodd C., Connell C. R., Heiner C. *et al* (1986). Fluorescence detection in automated DNA sequence analysis. Nature 321: 674-679.

Smith R., Healy E., Siddiqui S., Flanagan N., Steijlen P. M., Rosdahl I., and Jacques J. P. *et al.* (1998). Melanocotin 1 receptor variants in an Irish population. J Invest Dermatol 111: 119-122.

Soodyall H. and Jenkins T. (1998). Khoisan prehistory: The evidence of genes. Section 7: Khoisan biology. In: Bank A. (Ed.). The proceedings of the Khoisan identities and cultural heritage conference. Pages 374-382.

Spritz R. A., Strunk K. M., Hsieh C., Sekhon G. S., and Francke U. Homozygous tyrosinase gene mutation in an American black with tyrosinase-negative (type 1A) oculocutaneous albinism (1991). Am J Hum Genet 48: 318-324.

Spritz R. A. and Hearing V. J. (1995). Genetic disorders of pigmentation. Chapter 1. In: Advances in Hum Genet 22: 1-45.

Spritz R. A., Bailin T., Nicholls R. D., Lee S-T., Park S-K., Mascari M. J., and Butler M. G. (1997). Hypopigmentation in the Prader-Willi syndrome correlates with *P* gene deletion but not with haplotype of the hemizygous *P* allele. Am J Med Genet 71: 57-62.

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

Sturm R. A., Baker E., and Sutherland G. R. (1994). Assignment of the tyrosinase-related protein-2 gene (TYRP2) to human chromosome 13q31-q32 by fluorescence *in situ* hybridization: Extended synteny with mouse chromosome 14. Genomics 21: 293-296.

Sturm R. A., O'Sullivan B. J., Box N. F., Smith A. G., Smit S. E., Puttick E. R. J., Parsons P. G. *et al* (1995). Chromosomal structure of the human TYRP1 and TYRP2 loci and comparison of the tyrosinase-related protein gene family. Genomics 29: 24-34.

Sturm R. A., Box N. F., and Ramsay M. (1998). Human pigmentation: The difference is only skin deep. Bioessays 20: 712-721.

Suzuki I., Tada A., Ollmann M. M., Barsh G. S., Im S., Lamoreux M. L., Hearing V. J. *et al.* (1997). Agouti signaling protein inhibits melanogenesis and the response of human melanocytes to α-melanotropin. J Invest Dermatol 108: 838-842.

Tajima F. and Nei M. (1984). Estimation of evolutionary distance between nucleotide sequences. Mol Biol Evol 1: 269-285.

Taylor G. R. (1992). Polymerase chain reaction: Basic principles and automation. Chapter 1. In: McPherson M. J., Quirke P., and Taylor G. R. (Eds.) PCR- a practical approach. Oxford University Press. Pages 1-14.

Theophilus B., Latham T., Grabowski G. A., and Smith F. I (1989). Gaucher disease: molecular heterogeneity and phenotype-genotype correlation. Am J Hum Genet 45: 212-225.

Thody A. J., Higgins E. M., Wakamatsu K., Ito S., Burchill S. A., and Marks J. M. (1991). Pheomelanin as well as eumelanin is present in human epidermis. J Invest Dermatol 97: 340-344.

Tobias P. V. (1974). An anthropologist looks at malaria. S Afr Med J 48: 1124-1127.

Tobin D., Quinn A. G., Ito S., and Thody A. J. (1994). The presence of tyrosinase and related proteins in human epidermis and their relationship to melanin type. Pigment Cell Res 7: 204-209.

Tomita Y., Takeda A., Matsunaga J., Okinaga S., Shibahara S., and Tagami H. (1992). Molecular bases of tyrosinase-negative oculocutaneous albinism: a single base insertion or a missense point mutation in the tyrosinase gene. Pigment Cell Res (suppl 2): 96-100.

Tomita Y. (1994). The molecular genetics of albinism and piebaldism. Arch Dermatol 130: 355-358.

Tripathi R. K., Giebel L. B., Strunk K. M., and Spritz R. A. (1991). A polymorphism of the human tyrosinase gene that is associated with temperature-sensitive enzymatic activity. Gene Expression 1: 103-110.

Tripathi R. K., Hearing V. J., Urabe K., Aroca P., and Spritz R. A. (1992a). Mutational mapping of the catalytic activities of human tyrosinase. J Biol Chem 267: 23 707-23 712.

Tripathi R., Strunk K. M., Giebel L. B., Weleber R. G., and Spritz R. A. (1992b). Tyrosinase gene mutations in type I (tyrosinase-deficient) oculocutaneous albinism define two clusters of missense substitutions. Am J Med Genet 43: 865-871.

Tripathi R. K., Bundey S., Musarella M. A., Oroetto S., Strunk K. M., Holmes S. A. and Spritz R. A. (1993). Mutations of the tyrosinase gene in Indo-Pakistani patients with type I (tyrosinase deficient) oculocutaneous albinism(OCA). Am J Hum Genet 53: 1173-1179.

Uhler M., Herbert E., D'Eustachio P., and Ruddle F. D. (1983). The mouse genome contains two non-allelic pro-opiomelanocortin genes. J Biol Chem 258:9444-9453.

Vage D. I., Lu D., Klungland H., Lien S., Adalsteinsson S., and Cone R. D. (1997). A non-epistatic interaction of *agouti* and *extension* in the fow, *Vulpes vulpes*. Nature Genet 15: 311-315.

Valverde P., Healy E., Jackson I., Rees J. L., and Thody A. J. (1995). Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nature Genet 11: 328-330.

Valverde P., Healy E., Sikkink S., Halden F., Thody A. J., Carothers A., Jackson I. J. *et al.* (1996). The Asp84Glu variant of the melanocortin receptor (*MC1R*) is associated with melanoma. Hum Mol Genet 5: 1663-1666.

Van Dorp D. B., Van Haeringen N. J., and Glasius E. (1982). Evaluation of hairbulb incubation test and tyrosine assay in the classification of albinism. Ophthal Paediatrics and Genet 1: 189-200.

Vaughan D. and Asbury T. (1977). Growth, development and senescence. Eighth edition.

Walsh R. K. (1971). A distinctive pigment of the skin in New Guinae indigenes. Ann Hum Genet 34: 379-388.

Wilson B. D., Ollmann M. M., L. Kang., Stoffel M., Bell G. I., and Barsh G. S. (1995). Structure and function of ASP, the human homolog of the mouse agouti gene. Hum Mol Genet 4: 223-230.

Winder A. J., Wittbjer A., Rosengren E., and Rorsman H. (1993). The mouse brown (b) locus protein has a dopachrome tautomerase activity and is located in lysosomes in transfected fibroblasts. J Cell Sci 106: 153-166.

Winder A., Kobayashi T., Tsukamoto K., Urabe K., Aroca P., Kameyama K. and Hearing V. J. (1994). The tyrosinase gene family- interactions of melanogenic proteins to regulate melanogenesis. Cell Mol Res 40 (7-8): 613-626.

Witkop C. J., Quevedo W. C., and Fitzpatrick T. B. (1983). Albinism and other disorders of pigment metabolism. Chapter 15. In: Stanbury J. B., Wyngaarden J. B., Fredrickson D. S., Goldstein J. L., and Brown M. S. (Eds). The metabolic basis of inherited diseases. Fifth edition. McGraw-Hill Book Company. Pages 301-346.

Xu X., Thornwall M., Lundin L-G., and Chhajlani V. (1996). Val92Met variant of the melanocyte stimulating hormone receptor gene. Nature Genet 14:384.

Yang Y-K., Dickinson C., Haskell-Luevano C., and Grantz I. (1997). Molecular basis for the interaction of [Nle^4,D-Phe^7]melanocyte stimulating hormone with the human melanocortin-1 receptor (melanocyte α-MSH receptor). J Biol Chem 272: 23 000-23 010.

Yokoyama K., Yasumoto K-I., Suzuki H., and Shibahara S. (1994). Cloning of the human DOPchrome tautomerase/tyrosinase-related protein 2 gene and identification of two regulatory regions required for its pigment cell-specific expression. J Biol Chem 269: 27 080-27 087.

Internet Sources:
Skin cancer- the facts (updated 1998). Http://homepages.tcp.co.uk/~jas/skincancer/facts.html
National psoriasis Foundation (updated 1998). Http//www.psoriasis.org/whatis.html

Kits and manufacturers' protocols:
DIG users guide (Boeringer Mannheim).
SAP-Exonuclease protocol.
MDE protocol (FMC).
CDP-*Star*™ protocol (Boeringer Mannheim).
d-Rhodamine terminator cycle sequencing kit protocol (Perkin.Elmer)