Australian Government

Rural Industries Research and Development Corporation

Inheritance of White Colour in Alpacas

— Identifying the genes involved —

RIRDC Publication No. 11/074

RIRDC Innovation for rural Australia

Australian Government
Rural Industries Research and Development Corporation

Inheritance of White Colour in Alpacas

Identifying the genes involved

by Kylie Munyard

July 2011

RIRDC Publication No. 11/074
RIRDC Project No. PRJ-000060

ISBN 978-1-74254-258-4
ISSN 1440-6845

Inheritance of White Colour in Alpacas: Identifying the genes involved
Publication No. 11/074
Project No. PRJ-000060

Researcher Contact Details

Kylie Munyard
School of Biomedical Sciences
Curtin University
GPO Box U1987
Perth WA 6845

Phone: 08 9266 7519
Fax: 08 9266 2342
Email: K.Munyard@curtin.edu.au

In submitting this report, the researcher has agreed to RIRDC publishing this material in its edited form.

RIRDC Contact Details

Rural Industries Research and Development Corporation
Level 2, 15 National Circuit
BARTON ACT 2600

PO Box 4776
KINGSTON ACT 2604

Phone: 02 6271 4100
Fax: 02 6271 4199
Email: rirdc@rirdc.gov.au.
Web: http://www.rirdc.gov.au

Electronically published by RIRDC in July 2011
Print-on-demand by Union Offset Printing, Canberra at www.rirdc.gov.au
or phone 1300 634 313

Foreword

This report describes the research conducted as part of the alpaca colour genetics project called "Inheritance of White colour in Alpacas: identifying the genes involved". Three approaches were used (Mendelian, physical and genetic) in an attempt to unravel the mystery surrounding colour inheritance in alpacas.

This project has successfully identified key mutations in genes that lead to differences in fibre colour in alpacas. Other genes, which play a role in colour variation in other species, were cleared of involvement in colour variation in alpacas. Through extensive observational analysis a model for Mendelian inheritance of the major colours was developed. In combination, these findings provide breeders with information that allows them to make informed colour breeding choices. By adopting a genetically-grounded colour nomenclature system breeders will be able to avoid the ambiguity that currently exists in the industry.

This project was funded by RIRDC Core funds with some voluntary funding provided by the AAA. The AAA also provided extensive in-kind contributions through their members. AAFL kindly provided fibre samples for analysis. Additional funds, in-kind assistance and facilities were provided by Curtin University.

This report is an addition to RIRDC's diverse range of over 2000 research publications and it forms part of our Rare Natural Animal Fibres R&D program, which aims to foster the development of rare fibre industries.

Most of RIRDC's publications are available for viewing, free downloading or purchasing online at www.rirdc.gov.au. Purchases can also be made by phoning 1300 634 313.

Craig Burns
Managing Director
Rural Industries Research and Development Corporation

About the Author

Dr Kylie Munyard obtained her undergraduate degree in Biology, with Honours in Veterinary Biology, from Murdoch University in 1990. After working as a Clinical Cytogeneticist, she went on to complete a PhD ("The Ecology of Methanogens in the Rumen") at UWA/CSIRO in 2000.

Dr Munyard is currently a Senior Lecturer/Researcher in Molecular Biology at the School of Biomedical Sciences at Curtin University. Her research interests are South American Camelid genetics and conservation genetics of Australian quail. In 2006, with funding from the AAA, she worked with alpaca breeders to perform a survey of the incidence of staggers in Australian alpacas. A two-year project investigating microsatellite markers in alpacas (funded by Morris Animal Foundation) has also been completed. The results from both projects have been published in peer-reviewed scientific journals. In 2008 she was an invited speaker at two seminal events, the Inaugural World Alpaca Conference (Sydney) and the 1st International Workshop on Camelid Genetics.

Prior to commencing Academic duties in the school of Biomedical Sciences at Curtin University, Dr Munyard was a Postdoctoral Researcher at the Centre for High-throughput Agricultural Genetic Analysis (CHAGA), specialising in animal and instrumentation projects. Her projects within CHAGA included the invention of a DNA extraction kit, as well as provision of parentage assessments on cattle, and the development of an online database that allows fast and easy assignment of parentage to unknown animals. Two provisional patents and one full patent have arisen from that work.

Acknowledgments

The research reported herein would not have been possible without the unstinting, enthusiastic support of the members of the AAA. They provided advice, support, information and, most importantly, samples from their fascinating and precious animals. Special thanks go to Veterinarian Carolyn Emery for her assistance with sample collection throughout the project.

Many Students from Curtin's School of Biomedical Science have been involved in this project, and I thank them for their contributions. I would particularly like to acknowledge the extensive contributions of Ms Natasha Feeley and Mr Rhys Cransberg. Both Tash and Rhys received First Class Honours, and are now studying towards their PhD, on research directly related to this project.

Abbreviations

DNA	Deoxyribose nucleic acid
RNA	Ribose nucleic acid
AAA	Australian Alpaca Association Ltd
AAFL	Australian Alpaca Fleece Ltd
PCR	the polymerase chain reaction
bp	base pairs
Kb	kilobases
SNP	single nucleotide polymorphism
RNA-seq	massively parallel sequencing of expressed RNA

Contents

Tables

Figures

Executive Summary

What the report is about

This project has generated essential basic information about the Mendelian inheritance and molecular genetics of fibre colour in alpacas. Such basic science underpins successful animal breeding, and is information that is known about every other commercially produced livestock species. It is therefore important for the alpaca industry to increase the amount of knowledge about colour genetics.

Who is the report targeted at?

This report is targeted at alpaca breeders worldwide. Whilst the data has been collected from Australian alpaca herds, the results described herein are applicable to all alpacas, both huacaya and suri. Breeders will be able to use this information to increase the predictability of colour outcomes in their breeding programs.

Background

This project was designed to meet the growing need for knowledge about inheritance of colour in alpacas. Colour, also known as the pigmentation process, is important for many reasons, including camouflage, protection from UV radiation, commercial value and as a role in removing toxic by-products of metabolism. Modern alpaca breeders do not have enough information at their disposal to make informed breeding choices in relation to colour, due largely to the substantial loss of knowledge at the time of the Spanish conquest, and to subsequent lack of research into the field. In very recent times, molecular genetics tools have been made available to the alpaca research community that allow us to study pigmentation in alpacas much more easily and thoroughly. These tools include the alpaca genome sequence and massively parallel nucleic acid sequencing. However, it is important that these molecular tools are used in conjunction with the more traditional tools of Mendelian genetics and physical analysis, so that underlying biological processes are always in consideration when trying to answer questions of biology.

Aims/objectives

This project had four major objectives:

- To determine the Mendelian inheritance of white and other coloured fibre in alpacas.
- To determine the sequence of the coding regions of the three major colour-related genes agouti, melanocortin-1 receptor, and tyrosinase related protein, and to try to identify mutations that are responsible for the variety of colours in alpacas.
- To determine the alpaca mRNA expression profile of up to 20 genes known to be involved with pigmentation in other mammals.
- To determine the physical characteristics of the melanins present in alpaca fibre.

Methods used

A combination of methods was used during this research project. Visual inspection of animals was combined with breeding records to develop a model of Mendelian inheritance of colour in alpacas. Chemical analysis of fibre was used to determine the relative amount and type of melanins present in the fibre. Finally, molecular genetics techniques including DNA extraction, PCR, DNA sequencing and RNA sequencing were used to generate data about the genes controlling colour in alpacas.

Results/key findings

As a result of research conducted during this project alpaca breeders now have a comprehensive model of Mendelian inheritance of colour. The complete coding sequence of five major colour genes have been determined (*Mc1r*, *Asip*, *Tyrp1*, *Tyr*, *Matp*). Two of these genes (*Mc1r* and *Asip*) contain mutations that affect colour. RNA-seq was used to ascertain the expression pattern of over 5000 genes from alpaca skin. Forty one genes known to affect colour in other species had measurable expression in alpaca skin. Of these, 15 had different levels of expression in different colours. The differential expression of genes *Mart-1* and *mir22* could not be explained by reference to known pigmentation processes, and so are strong candidates for further study into alpaca pigmentation. In addition, variable expression levels were detected in 60 other genes that have not previously been linked to colour variation in other species. Finally, we used physical analysis to examine the type and amount of pigment present in 11 different colour classes of alpaca fibre. These analyses support the hypothesis that eumelanic brown does not exist in alpacas, and that brown alpacas are genetically red (or yellow). Breeders will be able to use all this information to make more predictable colour breeding outcomes, and the Industry will benefit as a result of increased production of the desired fibre colours.

Implications for relevant stakeholders

The research outcomes presented in this report have implications for the alpaca industry, alpaca breeders, policy makers and for Camelid scientists. The Industry, through its breeders, now has valuable information about the way that colour is determined at a genetic level in alpacas. This knowledge will enable breeders to make informed breeding decisions that will better enable them to meet Industry demands for specific fibre colours. Policy makers can use this information to tailor programs such as registration and show rules to more closely align with genetics underlying colour, rather than on visual examination alone. Camelid researchers will be able to use the information presented herein to continue to advance the state of knowledge of alpaca and llama colour genetics.

Recommendations

The recommendations of this report are targeted at alpaca breeders and policy makers in Australia and worldwide. It is imperative that a common, genetically based nomenclature system be introduced into the alpaca industry. Without a common language for colour, the current mis-identification and confusion will continue. In addition, the colour of all cria should be recorded (including pattern) preferably by photography, so that any age-related changes can be accounted for in breeding programs, especially when animals change ownership. It is the unusual or unexpected colour outcomes that hold the key to solving the complete alpaca colour inheritance riddle.

The chemical test for melanin content and the tests for DNA mutations associated with some colours should be used by breeders who want to have more certainty in their colour outcomes. Further research is needed to clarify the physical characteristics of pink-skinned white, black skinned white, and blue-eyed white.

Despite the identification of mutations relating to base colour, there is still no genetic explanation for the variation in intensity of those colours within a given genotype. The desired intensity can be selected for in a breeding program; however this will take many years to have a noticeable effect. In addition, it is not clear if the variation is due to genetics, environment or a combination of both. Further phenotypic and molecular research is needed to elucidate the cause of the variation.

Introduction

At the August 2006 AAA National conference, delegates were told that there was a significant unfilled demand for white alpaca fibre. Data was also presented showing that while the percentage of white alpacas in the Australian herd had risen in the past 10 years, it was still well below 40%. Breeding records from the AAA herd books show that there is no clear path to the guaranteed production of white animals. For example, only 70% of white to white matings produce white offspring. It was clear that a better understanding of the colour genetics of alpacas was required to help the industry to be able to meet the demand for white fibre. At present, there is no universally accepted theory that can explain how white, coloured and patterned alpacas occur. Some progress has been made (e.g. Frank *et al.* 2005). However, these theories have not been fully validated. Therefore, this project was designed to combine Mendelian, physical and molecular methods in an attempt to elucidate the basics of colour inheritance in alpacas.

Why we should study colour genetics

The answer to the question "why study colour in alpacas?" might seem like a foregone conclusion. And it is; the colour of an alpaca fleece is important for many obvious reasons. However, the answer should not stop there. Some of the other reasons why alpaca colour should be studied include the role of pigment in protection from UV radiation, in vitamin D dose, and as a role in removing toxic by-products of metabolism. Colour in mammals has been widely studied and is fairly well understood. This is because the pigmentation process is a very good model for many cellular processes, and because changes in colour are easy to see, and are relatively easy to understand.

There is currently no definitive model which explains the mechanisms that govern fibre colour inheritance in alpacas and therefore breeding programs are not effective in selecting for colour (McGregor 2006). Current breeding programs utilise the phenotype of the parents to predict breeding outcomes (Australian Alpaca Fleece Limited 2004, Paul 2006). However, phenotype can be a poor indicator of genotype because of the complex nature of pigment regulation in mammals (Hoekstra 2006, Rees 2003). Fibre colour phenotype is the outcome of the combined effect of over 200 genes (Furumura *et al.* 1996, Sponenberg 2001). The ability to determine fibre colour inheritance in alpacas would be highly valuable to the industry. It would facilitate more effective colour selection and the creation of a herd that is better suited to market demands.

Biology of Colour in Mammals: The Components

In general, pigmentation in mammals is a highly conserved process although subtle differences between species do exist. That is, the way that colour is generated and controlled is the same in all mammals. The differences that we see between species and within species are caused by variation within the genes that control colour. Some species have only one version (allele) of a gene, while others may have many alleles. A good example of this is the *Mc1r* gene. In humans there are many different versions, while in Golden Retrievers there is only one version.

Pigmentation is a complex process that begins in the early stages of the development of an embryo. In summary, neural crest cells differentiate into melanoblasts which then migrate along the back and down the sides of the body. As the melanoblasts migrate, they differentiate into melanocytes. Melanocytes produce internal organelles called melanosomes in which pigment granules called melanin are produced in a process called melanogenesis. The mature melanosomes are exported from the melanocyte into specialist skin and hair cells called keratinocytes, to produce the phenotype, the visual colour that we can see.

Melanocytes

A good understanding of the pigmentation process is essential in order to study how a phenotype is created. There are two types of neural crest-derived cells; early and late migrating. The early migrating neural crest cells become the peripheral nervous system, craniofacial bones, the heart and the adrenal medulla. The later migrating neural crest cells become melanoblasts, which eventually differentiate into melanocytes as they migrate through the body. There are four basic types of melanocytes, ear, eye, central nervous system, and skin/hair. The whole process of the development of the melanoblasts is usually completed during embryogenesis. Melanocytes are located slightly differently in skin and in hairy parts of the body (Figure 1.1). Melanocytes have protuberances called dendrites that are interspersed between the neighbouring keratinocytes (Figure 1.2). Skin and hair become coloured when mature melanosomes migrate from the melanocyte dendrites into the keratinocytes.

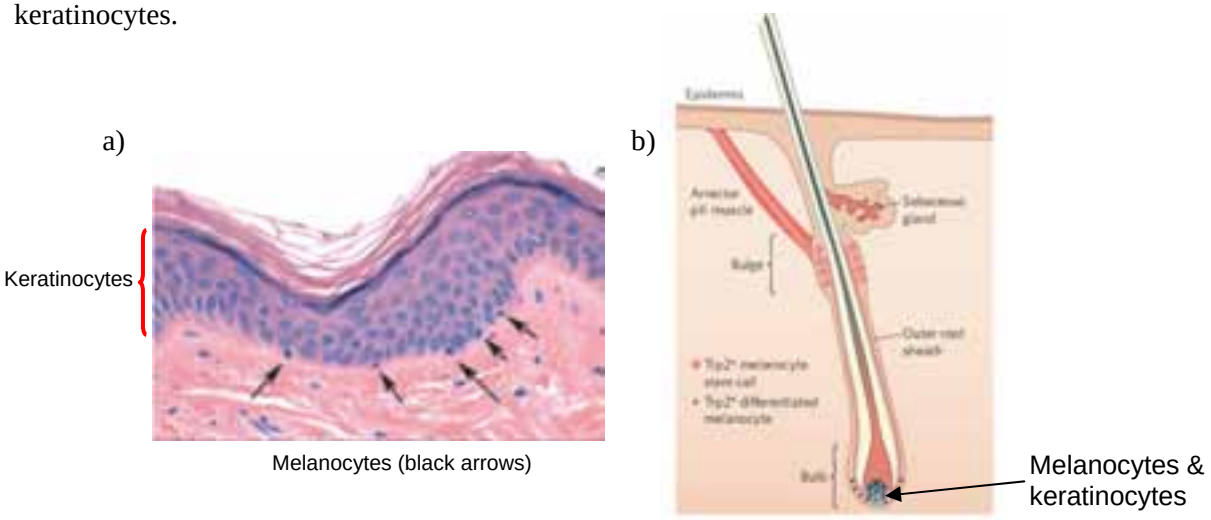

Figure 1.1: Location of skin (a) and hair (b) melanocytes (adapted from Lin & Fisher 2007).

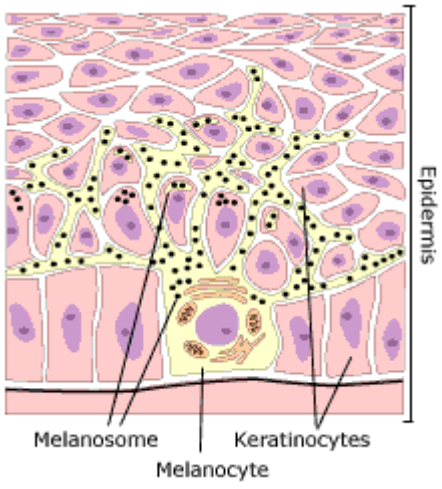

Figure 1.2: A melanocyte with dendritic processes (from: http://www.hominides.com/html/dossiers/race.html)

Melanosomes

Melanosomes are the site of melanin synthesis. Melanosomes are highly specialised organelles that are generated in the cytoplasm of melanocytes. Melanosome development follows four distinct stages, I to IV. During Stage I the melanosome is formed from a piece of the endoplasmic reticulum. A protein called Pmel17 is then transported into the melanosome and is cleaved into its active form. The Pmel17 protein forms a fibrillar structure within the melanosome during Stage II. This acts as a scaffold that stabilises the melanosome. During stage II the enzymes required for melanin synthesis are transported into the melanosome. Stage III marks the beginning of melanogenesis, and Stage IV is reached when the melanosome is full of pigment and is ready for transport into the keratinocytes.

Melanin Synthesis

Melanin synthesis (melanogenesis) occurs entirely inside the melanosome. In mammals there are only two types of melanins, eumelanin and pheomelanin. Eumelanin is dark in colour, either black or brown, while normal pheomelanin is yellow to red in colour. The precursor molecule for both types of melanin is the amino acid tyrosine. Tyrosine is converted through a series of reactions and intermediate products into either pheomelanin or eumelanin (Figure 1.3). Melanosomes that contain eumelanin (eumelanosomes) are ellipsoidal and smaller than pheomelanosomes (pheomelanin-containing melanosomes) which are spherical and about 0.7µm in diameter.

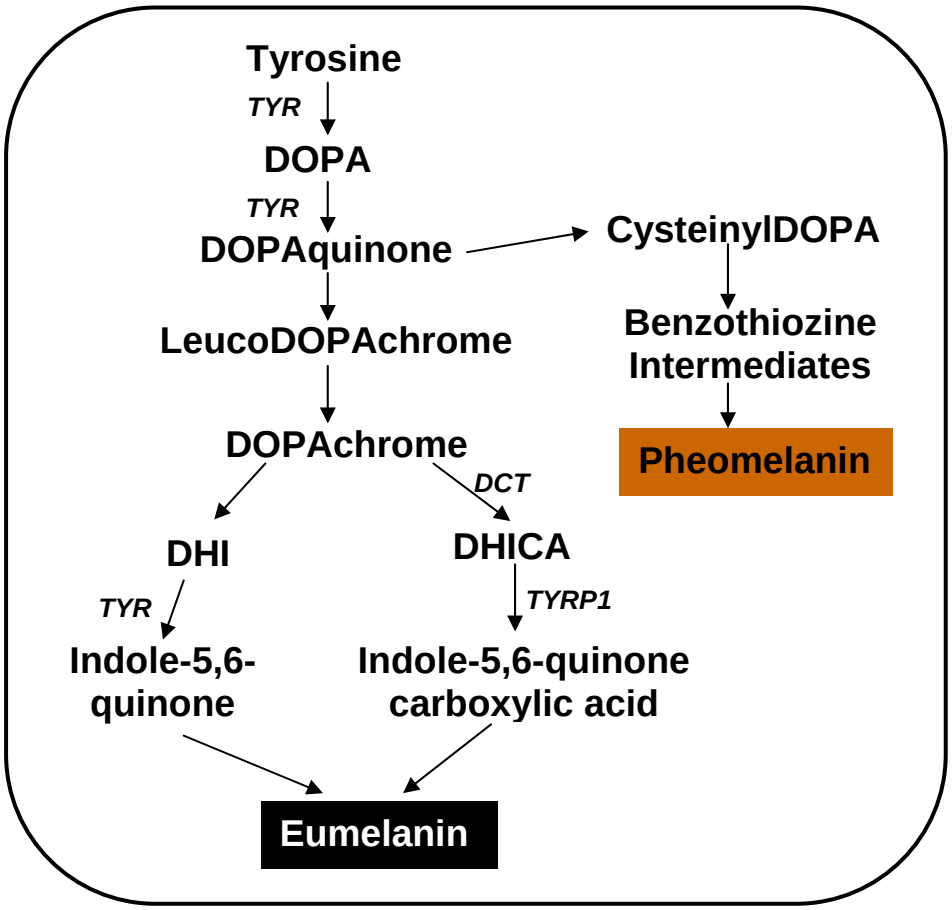

Figure 1.3: The melanin biosynthesis pathway.

Melanin synthesis also results in toxic by-products such as hydrogen peroxide and quinone intermediates. It is hypothesised that one reason for the sequestration of melanin synthesis into melanosomes is to protect the cell from these toxic compounds.

Control of Pigmentation in Mammals

The final phenotype of a mammal is affected by the enzymes and other proteins involved in all of the stages of pigmentation; from the development of the neural crest cells to the process of transfer of melanosomes into keratinocytes. Although we know a lot about the process we do not know everything about every step in the process. A very good example of this is the recent discovery that the β-defensin gene, previously believed to only play a role in the immune system, controls colour in dogs (Candille *et al.* 2007).

There are over 150 cloned colour loci (Lamoreux *et al.* 2010) some of which are shown in Figure 1.4. Each acts at a different part of the pigmentation process. Generally, they can be grouped into subcellular, cellular, tissue and environment genes. For example, a white forehead, chest and feet is a common pattern in many species, and is caused when the melanocytes don't quite reach the extremities of the body. This "white spotting" is caused when melanocytes do not complete their normal migration, and are confined to certain parts of the body, i.e. it is tissue specific. The transcription regulator microphthalmia-associated transcription factor (*Mitf*) controls the extent of melanocyte migration in some species, leading to various patterns of white on a solid colour background. Other genes act at the sub-cellular level. For example the *Silv* gene produces the protein Pmel17 which is involved in the merle pattern seen in many dog breeds. Other genes control the export of melanosomes into keratinocytes.

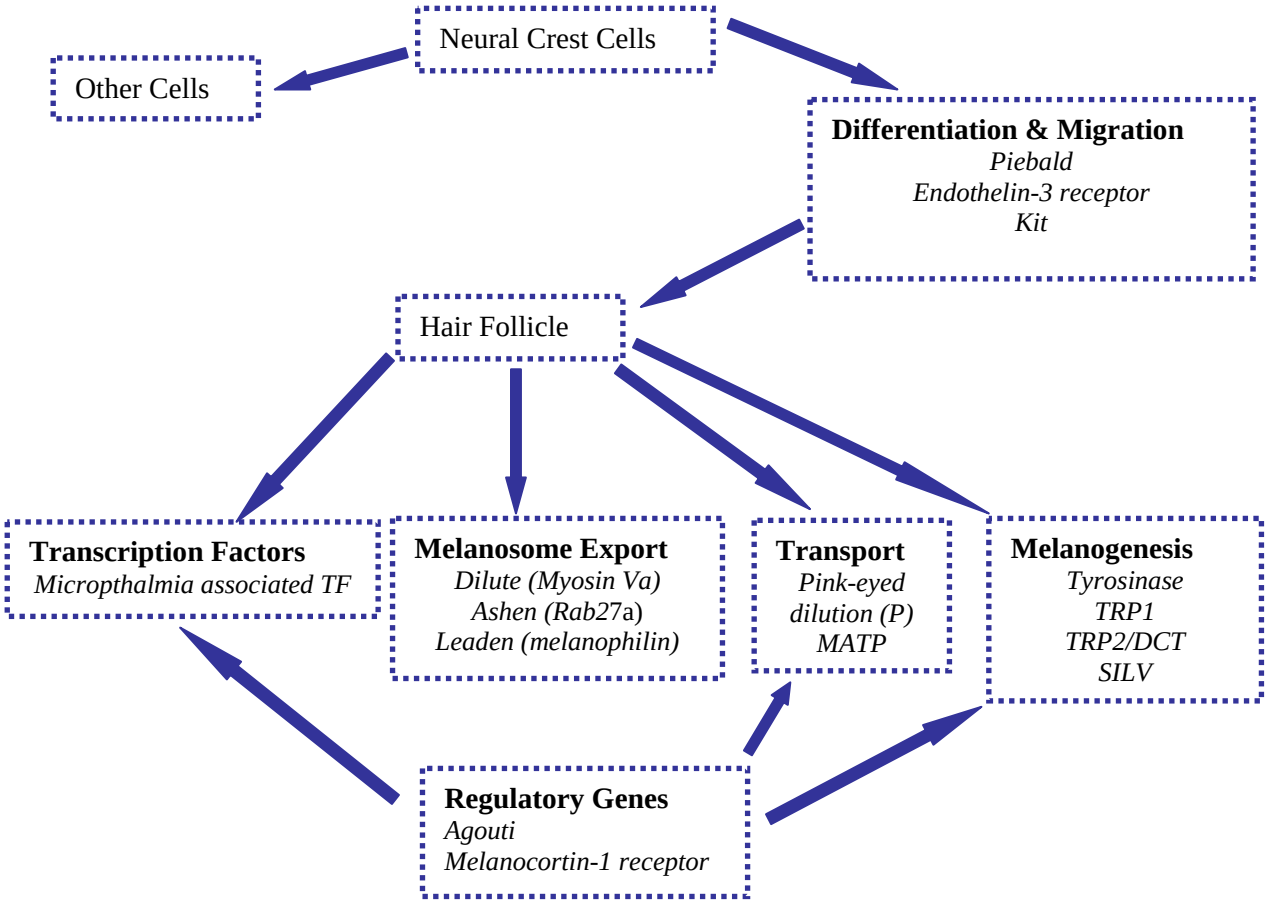

Figure 1.4: Examples of pigmentation genes and their point of action in the pigmentation pathway.

There is a three step process by which the mature melanosome is exported, and a mutation in any of the three genes in this process, myosinVa, Rab27a and melanophilin, leads to a failure of export, and therefore no (or reduced) pigment in the hair. When MyosinVa is faulty the reduction in the number of melanosomes that are exported leads to the well-known "dilute" phenotype where a normally black animal is slate grey.

The two key proteins controlling the expression of eumelanin and pheomelanin in the pigmentation pathway are encoded by the Agouti (*Asip*) and Melanocortin-1 Receptor (*Mc1r*) genes. Mc1r (the protein product of *Mc1r*) winds through the cell membrane of melanocytes, and agouti signalling protein (Asip, the protein product of *Asip*) is able to bind to it and switch melanin production to pheomelanin (Figure 1.5a). If Asip is not present, then alpha-melanocyte stimulating hormone (α-Msh) will bind to Mc1r instead, and only eumelanin will be produced (Figure 1.5b). In a normal cell, all three proteins are present and functional, leading to the production of both pheomelanin and eumelanin. When either gene produces a protein that is non-functional or not completely normal, changes in pigmentation occur. If *Mc1r* is mutated so that it is unable to be bound by α-Msh, then only pheomelanin will be produced. Examples of this type of mutation include Golden Retrievers and Labradors, chestnut coloured horses, and some red-haired humans. Sometimes *Mc1r* is mutated in such a way that it is "on" all the time. This is called an activation mutation, and leads to only eumelanin being produced regardless of whether other ligands are present or not. This mutation is known as the "dominant black" mutation and has been characterised in sheep.

a) **b)**

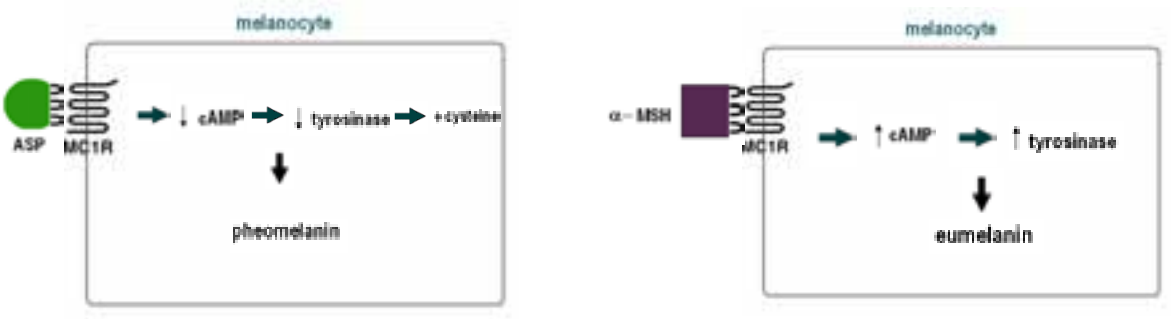

Figure 1.5: Interaction of Agouti signalling protein (ASP) and melanocyte stimulating hormone (α-MSH) with the melanocortin-1 receptor (MC1R; From: Feeley 2007). cAMP = cyclic AMP.

An understanding of the basic process of mammalian pigmentation allows the creation of hypotheses about how pigmentation is controlled in alpacas. This process begins with a comparison of the phenotype of alpacas with the phenotypes of other species (e.g. Figure 1.6). The similarities between the colour and pattern of alpaca and (for example) horses and dogs is striking, and leads to hypotheses about which genes may have variant alleles in alpacas.

In horses the chestnut colour is caused by a mutation in the *Mc1r* gene which causes a premature stop codon, resulting in a non-functional Mc1r protein. *Mc1r* is a small gene, with a single coding exon of 954 bases, and is therefore a good first target for DNA studies on alpaca colour.

The agouti gene controls the extent of pale versus dark melanin in the coat. Bay horses, Doberman dogs, some cattle colours and many mouse colours are caused by variants of the agouti gene (Figure 6). Agouti has over 30 variants in the mouse, and at least three are thought to exist in the dog. The most dominant agouti alleles lead to a pale phenotype, while the most recessive alleles cause black. The alleles in between these extremes cause combinations of pale and dark melanin that leads to the many patterns that exist.

Tyrp1 produces an enzyme that acts in the melanin synthesis pathway (Figure 1.3) and is known colloquially as the "brown" locus. A mutated *Tyrp1* causes brown eumelanin instead of black. The recessive brown variant of *Tyrp1* is found in dogs and mice, but does not exist in horses. It is possible that no brown alleles exist in alpacas because breeding outcomes suggest that brown in alpacas is dominant, rather than recessive. It is also possible that the same phenotype (i.e. brown) can be caused by two different mechanisms in alpacas and that that both recessive brown and dominant brown (probably occurring at the agouti locus) exist.

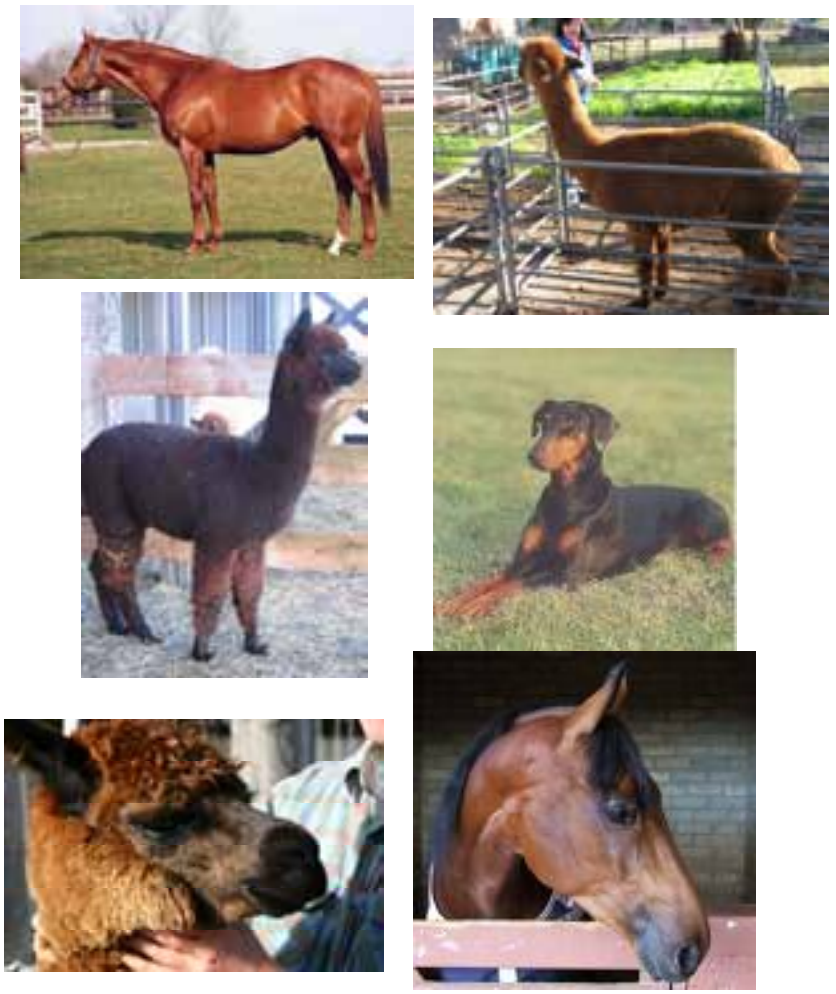

Figure 1.6: Comparisons of colour and pattern in different species.

Two other genes, tyrosinase (*Tyr*) and membrane associated transporter protein (*Matp*), may also be involved in colour variation in alpacas. *Tyr* is the gene that produces the enzyme tyrosinase, which is the key enzyme in the melanogenesis pathway (Figure 1.3). In species such as mice, rabbits and cats, variants of *Tyr* are known to produce a dilution of colour, including albinism (Lamoreux *et al.* 2010). *Matp* is the gene that is mutated in palomino and buckskin horses (Mariat *et al.* 2003). It has been hypothesised that some white alpacas are the equivalent of cremello horses. That is, they are chestnut with two doses of the dilution gene causing a massive reduction in the intensity of colour so that they appear white.

The following genes are hypothesized to have multiple variants in alpacas, and thereby cause colour variation: Agouti (A), Mc1r (E), Roan (Rn), Grey (M), Tuxedo (T), Piebald (P), Appaloosa. These genes can be broken into two groups, those that affect the base colour of the animal (*Mc1r* and Agouti) and those that cause alterations to the base colour (grey, roan, tuxedo, piebald and appaloosa). Other genes are almost certainly affecting colour, but these have not yet been identified, e.g. the gene(s) that

cause variation of intensity of colour within a genotype, such as the difference between blue-black and warm black, or dark skinned white and dark skinned pale fawn.

One of the problems with describing alpaca colour is the wealth of different words used to describe very similar phenotypes. Also, the colour of a shorn fleece can be different to the animal itself, e.g. a black roan and a silvergrey would both have their fibre classed as grey by AAFL. Therefore, photos are invaluable to show clearly what is meant by each description.

Physical Analysis of Colour

The large variety of colours present in all mammals is caused by the amount, placement and relative proportion of the two melanin pigments present in the fibre and skin. It can therefore be very difficult to objectively determine the colour of an animal by visual inspection alone. Accurate phenotype assignment is vitally important for correct comparisons between animals, for both Mendelian and molecular studies. It became apparent early in the project that an objective way to determine colour was essential. Handheld spectrophotometers have been used to generate colour spectra, however we felt that that type of analysis did not provide enough information for our purposes. A chemical method first used by Ozeki and colleagues (1996) was used to determine the total melanin content and proportion of eumelanin versus pheomelanin in alpaca fibre.

Molecular Genetics Analysis of Colour

There has been limited research into the genetic basis of colour in alpacas (Paul 2006). More research has been directed into fibre quality due to its financial significance (Frank et al. 2006; McGregor 2006). Prior to the start of this project, no molecular studies of alpaca colour genes had been published. Results from this project have been published in three papers, on *Mc1r* (Feeley & Munyard 2009), *Matp* & *Tyrp1* (Cransberg and Munyard 2011) and ASIP (Feeley *et al.* 2011). Since 2008, a group from the US have published a paper on *Mc1r* (Powell *et al.* 2008), and a conference publication from an Italian group describes work on *Mc1r* and *Asip* (Chandramohan *et al.* 2010). However, neither the US or Italian groups associated mutations with particular colour variants. Described in this report are the associations made between colour and DNA mutations as part of this project.

Objectives

This project had four major objectives:

- To determine the Mendelian inheritance of white and other coloured fibre in alpacas.
- To determine the sequence of the coding regions of the three major colour-related genes agouti, melanocortin-1 receptor, and tyrosinase related protein, and to try to identify mutations that are responsible for the variety of colours in alpacas.
- To determine the alpaca mRNA expression profile of up to 20 genes known to be involved with pigmentation in other mammals.
- To determine the physical characteristics of the melanins present in alpaca fibre.

Methodology

Animals

All research for this project was carried out with Curtin University Animal Ethics approval (R01-2009 and N43-09). Alpacas from Western Australia, Victoria, New South Wales and South Australia were included in this study. Both Suri and Huacaya, male and female were collected, and have been treated as one group for the purposes of colour analysis. Each animal was given a unique laboratory number so that it was de-identified for reporting purposes. At the time of collection digital photographs were taken to record the colour of the body, face, neck, feet and legs. A visual assessment of colour phenotype was performed, and recorded. Any patterns or unusual marking, and any information about age-related changes in colour were also recorded, as was a colour pedigree, if the breeder/ owner had access to that information.

Sampling

Three types of samples were collected from alpacas, blood, skin, and fibre. Approximately 5 mL of blood was collected into K_3EDTA tubes and stored at -20°C. Disposable sterile dermal punches (3 mm diameter) were used to collect skin samples from the shoulder region. Skin samples were stored in RNALater (Invitrogen) at -80°C. Fibre samples were collected from the body, and any other areas of differing colour (if possible). Approximately 200mg of fibre was collected and stored in air-tight plastic bags at room temperature.

Melanin analysis

The amount and type of melanin present in fibre samples was measured according to the method of Ozeki *et al.* (1996). Fibre from 55 samples supplied by AAFL, and 91 samples collected at a show were analysed. Briefly, approximately 100mg of fibre was washed in ethanol, air-dried, then ground in liquid nitrogen to form a fine powder. Various amounts of fibre, depending on colour, were weighed accurately and added to a 2mL screw-capped cryotube; White – 20mg, Light Fawn – 20mg, Fawn – 15mg, Light Brown – 15mg, Brown – 10mg, Dark Brown – 5mg, Grey – 5mg, or Black – 2mg. Nine hundred microlitres (µL) of Soluene-350 and 100µL H_20 were added, and mixed well using a vortex mixer. The samples were incubated at 99ºC for 45 minutes, with vortex mixing every 15 minutes. The samples were centrifuged at 3500 rpm for 5 minutes, transferred to quartz cuvettes and the absorbance (Abs.) at 500nm and 650nm was measured. The relative amount of melanin in a sample was determined by dividing the Abs. 500nm by the weight in mg, to give relative melanin/ mg fibre. The type of melanin present was determined via the ratio of Abs. 650nm divided by the Abs. 500nm. A ratio of greater than 0.25 indicates that the sample contains mostly eumelanin, a ratio below 0.15 indicates that the predominant melanin is pheomelanin, and an intermediate ratio indicates that both types of melanin are equally abundant.

DNA Extraction

DNA was extracted from alpaca blood using the salt precipitation method described by Miller *et al.* (1988). Where this method failed to yield sufficient quantities of DNA, or where the quantity of blood obtained was less than 5 mL, the DNeasy blood and tissue DNA extraction kit (Qiagen) or the Axyprep Genomic DNA Extraction kit (Axygen) was used, according to the manufacturer's instructions.

PCR

Genomic DNA (50-100ng) was used as template for PCR using 2 µM of forward primer, 2µM reverse primer, 1× Polymerisation buffer (67 mmol/L Tris.HCl (pH 8.8), 16.6 mmol/L $[NH_4]_2SO_4$, 0.45% v/v Triton X-100, 0.2 mg/mL gelatin; Fisher biotec), 2mM $MgCl_2$ (Fisher biotec) and 0.75u BIOTAQ

polymerase (Bioline) in a 10µL reaction. Optimized amplification conditions were: 95°C for 2 min; 30-40 cycles of 95°C for 20s, 55 - 65°C for 30s and 72°C for 1 min; then 72°C for 5 min. Details of primers and annealing temperatures for *Mc1r*, *Matp*, *Tyr* and *Asip* are given in Feeley and Munyard (2009), Cransberg and Munyard (2010) and Feeley *et al.* (2011), details of *Tyrp1* primers are shown in Table 1.1. Except for *Mc1r*, primers for gene exons were designed based on the alpaca assembly available on the Ensembl database (http://www.ensembl.org/index.html) and were designed to hybridise approximately 100bp outside of the predicted splice sites for exons, thereby amplifying the complete coding region of each gene. Amplified DNA was analysed by electrophoresis in 1.5% (w/v) agarose gels in TAE buffer, stained with ethidium bromide and visualised by UV transillumination.

Table 2.1: Primers designed for amplification of *Tyrp1* exons

Primer	Sequence	Product size (bp)	Temp (°C)
Ex2F	5′ CGGCTACATGGATTGACTTCC 3′	850	60
Ex2R	5′ TTCACTTTGAGGTGGGTTGG 3′		
Ex3F	5′ TGAAATTGCTTGGTCAGTGC 3′	525	61
Ex3R	5′ GCTCATCTCTACCCACGCTC 3′		
Ex4F	5′ AAGCCAAGCAAAGGGAGAAT 3′	560	65
Ex4R	5′ AGTGGGTTTTGGTGTGAAGC 3′		
Ex5F	5′ CCACATTACCTCAGGCAAGC 3′	2500	62
Ex5R	5′ ATGACCAGTGATGGGAGA 3′		
Ex5IF	5′ AATCACAGAAGTTGGACATGG 3′		60
Ex5IR	5′ ATCAATCTGGCATTCAAAGGT 3′		
Ex6F	5′ TGTTGAGCCTGCAAAAA 3′	304	55
Ex6R	5′ TGTTTCCCAATATCATCACTGT 3′		
Ex7F	5′ TTTTGGGTCACCTTCAGAACA 3′	287	57
Ex7R	5′ GGGTAACACATTTGCTTTTGG 3′		
Ex8F	5′ TTTGCTCTCATTTCCTTTTTCA 3′	306	57
Ex8R	5′ AGCTTTTAATTCCAACCTGTGC 3′		

DNA Sequencing

PCR products were amplified from genomic DNA in five separate 10µL reactions and pooled prior to sequencing. Pooled PCR products were purified using the AxyPrep PCR Cleanup Kit (Axygen, Union City, CA, USA). Sequencing was performed using the ABI Big Dye Terminator® system at either Murdoch University, Perth, using a 48-capillary ABI 3730 DNA analyser or at Macrogen, Korea, using 96-capillary ABI 3730XL DNA analysers.

Both Vector NTI software (Invitrogen) and Geneious 4.8.3 (Biomatters) were used to assess and visualise sequences, assemble contigs and perform alignments. Splice sites were determined using SpliceView (http://zeus2.itb.cnr.it/~webgene/wwwspliceview.html) coupled with known bovine and human exons. Complete coding sequences for each animal were compiled using Geneious 4.8.3 **(**Biomatters), and compared with genes and proteins from other species by GenBank NCBI BLASTn and BLASTx protocols (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

RNA Extraction & Sequencing

RNA was extracted from tissue samples by homogenisation in Trizol (Invitrogen) according to the manufacturer's protocols. Briefly, each tissue sample (<100 mg in weight) was homogenised in 1ml of Trizol using the Fastprep system (Thermo Life Sciences) using lysing matrix D beads and a setting of 4.0 m/s for 30 seconds. After room temperature incubation for 5 minutes, 200 µl of 2-bromochlorophenol was added and the tubes mixed. The aqueous phase was isolated by centrifugation at 12,000 RCF, for 15 minutes and 500 µl aspirated to a fresh RNase free tube. Isopropanol, 500 µl

was added to the isolated aqueous phase to precipitate the RNA which was collected by centrifugation at 12,000 RCF for 30 minutes. The RNA pellet was washed once in 75% ethanol before resuspension in 100 µl of RLT buffer (Qiagen RNAeasy kit). The RNAeasy kit was used to further purify the RNA according to the manufacturer's protocol, collecting the RNA onto the spin filter, washing and extracting the clean RNA in 50 µl of Tris HCl pH 8.0.

The quality of the isolated RNA was assessed using RNA pico 6000 kit on the Agilent Bioanalyser. Only RNA with an RNA integrity value >7.0 was selected as suitable for sequencing. RNA was quantified using absorbance at 260nm on a Nanodrop spectrophotometer. Three individual pools of RNA were prepared by pooling equal amounts of RNA from all samples within each skin colour group.

Sequencing libraries were prepared from each of the three RNA pools using the Illumina TrueSeq RNA sequencing library preparation kit according to the manufacturer's protocol. Briefly 5 µg of the prepared total RNA was used in each reaction, the mRNA was captured using oligo d(T) magnetic beads. The purified mRNA was fragmented to an average size of 200bp, reverse transcribed using random primers and a double stranded cDNA created. The double stranded cDNA was blunt ended and A tailed before ligation of TrueSeq sequencing linkers, different primers were used for each pool so that sequencing could be used to identify the origin of the sequence generated. The cDNA with linkers was purified from unbound linkers and amplified with 15 rounds of PCR to produce sufficient DNA for sequencing. Primer dimers were removed to clean up the amplified DNA before QC analysis and quantification.

Each sequencing library was analysed using the Agilent Bioanalyser DNA 1000 electrophoresis chip for DNA size and purity from non specific products. The quantity of DNA in each library was assessed using the Kapa Biosystems Library Quant kit.

The sequencing was performed on an Illumina Genome Analyser IIx with a standard paired end sequencing flow cell. The flow cell was prepared using an Illumina Cluster station and a paired end cluster generation kit. The three sequencing libraries were pooled in equimolar concentrations to a final concentration of 8 pM DNA for loading onto the flow cell. The libraries were loaded into a single lane on the flow cell, with the other lanes being occupied by different experiments. The flow cell was sequenced using the Illumina GAIIx using the sequencing control program SCS2.6 and RTA 1.6 for analysis. The recipe or the reaction was a 54 cycle multiplexed paired end run.

On completion of the run the data was analysed. Sequence data was de-multiplexed, separating reads according to their TruSeq index, and then converted into fastq format; both steps using Illumina software. Quality statistics were calculated and quality and nucleotide distribution plots were created using scripts from the fastx toolkit. These were used to determine that the sequencing was successful. CUFFLINKS v0.8.0 (Trapnell *et al.* 2010) was used to compare transcriptome sequences from different coloured animals.

Statistics

Statistical analyses were performed using the JMP® 7 statistical package (SAS institute Inc. 2007). The Pearson χ^2 test was used to investigate if an association existed between fibre and/or skin colour and mutations. Statistical significance was accepted at $p \leq 0.05$.

Results & Discussion

Mendelian Inheritance of Colour

The records collected throughout this project support the hypothesis that *Asip* and *Mc1r* are the key genes controlling colour in alpacas. There is evidence for four *Asip* alleles and two *Mc1r* alleles. A combination of these alleles plus the proposed pattern gene alleles is sufficient to explain all colour variation (including pedigree history) seen throughout this study. Other alleles may exist in the source population in South America.

ASIP (A)

Agouti is one of the two base colour genes. We hypothesize that four alleles exist for the ASIP in alpacas. In order of dominance these are:

- A = white to fawn
- A^b = brown with dark points (e.g. bay in horses)
- a^t = black & tan (e.g. Doberman markings)
- a = black (both warm & blue black)

In this allelic series, A is completely dominant to a^t and a, and incompletely dominant to A^b. A^b is incompletely dominant over a^t and a.

There are 10 possible combinations (genotypes) of these four alleles

- AA = white to fawn
- Aa^t = white to fawn
- Aa = white to fawn
- AA^b = fawn/ light brown
- A^bA^b = brown with black points
- A^ba^t = dark brown body with black points
- A^ba = black brown
- a^ta^t = black & tan
- a^ta = dark brown
- aa = black

Agouti alleles are probably responsible for most phenotypes in alpacas. All animals expressing an Agouti phenotype have black skin regardless of the colour of their fibre. All animals with black skin are probably expressing an agouti phenotype, but that phenotype may be obscured or altered by one of the pattern genes. The difficulty in assigning an accurate genotype (in the absence of DNA tests and/or breeding outcomes) is that there is a range of phenotypes for each genotype. This phenomenon has been proven in horses, where DNA tests are available for both the "a" and "a^t" alleles. In some cases horses proven by DNA tests to be aa and a^ta^t can be very difficult to tell apart visually, while in other cases they can be very easily distinguished. Alpacas with AA, Aa^t or Aa genotypes will not have darker points ("points" refers to the muzzle, ears, around eyes, lower legs, sometimes tail; Figure 3.1).

Figure 3.1: Agouti genotypes AA, Aa^t & Aa. These three different genotypes all look the same, and can have white to fawn fibre colour with black skin.

AA^b animals will have darker fibre than other A- genotypes (Figure 3.2). With A^{b}a and A^{b}a^t heterozygotes (Figure 3.3), the colour of the brown on the body darkens compared with the A^{b}A^b homozygote (Figure 3.4), and black points may be more extensive. A^{b}A^b, A^{b}a^t & A^{b}a animals will have brown to dark brown fibre at the skin.

Figure 3.2: The AAb genotype has darker fibre than the AA, Aat or Aa genotypes, but still has black skin and an absence of black points.

Black & Tan animals have the characteristic presence of pheomelanic pigment on the ventral surface (e.g. chin, chest, armpits, belly, groin, and under the tail; see Figure 3.5). The extent of pheomelanic pigment varies from extensive to minimal. In black animals (i.e. aa), regardless of whether blue-black or warm black, the whole body will be the same colour, without obvious darkening of the points

(Figure 3.6). Black animals will have black fibre at the skin, but may have paler fibre tips and topknots due to weathering. Heterozygotes of these two alleles (i.e. $a^t a$) will have minimal ventral pheomelanin, but may also appear to be a less intense black overall.

Figure 3.3: $A^b a^t$ & $A^b a$ genotypes. The black points are still clearly present, and the colour of the body is darker, sometimes with the appearance of a black overlay (image on right) and is often described as mahogany.

Figure 3.4: The $A^b A^b$ genotype has clear black pigment on the muzzle, around the eyes, on the ears and on the lower legs. The body colour is bright red to dull red.

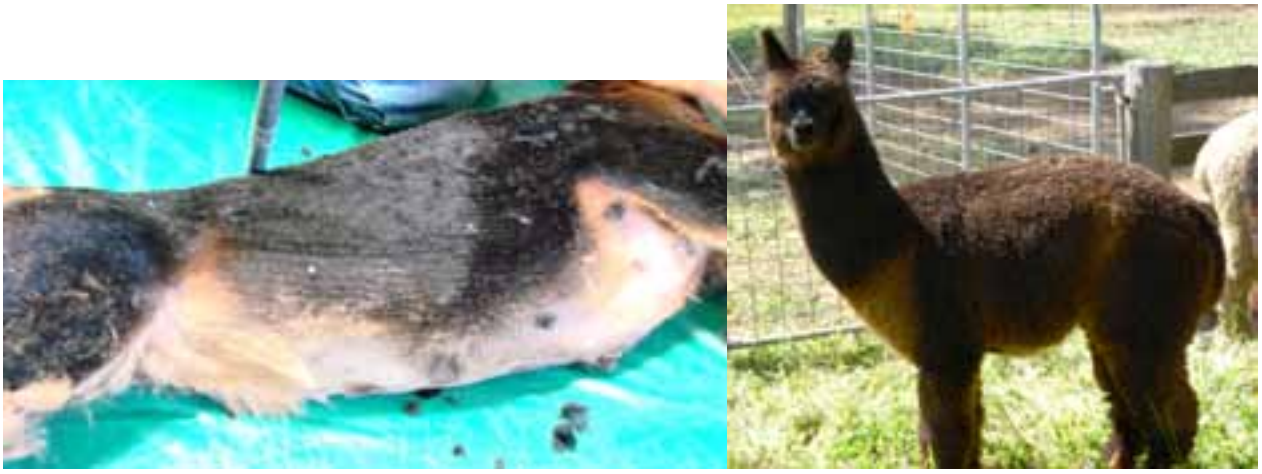

Figure 3.5: Black & Tan phenotypes of the genotype $a^t a^t$. Note the characteristic presence of pheomelanic fibre on the ventral surfaces, including the chin, chest, armpits, groin and under the tail.

Figure 3.6: Black phenotypes, with genotype aa. The effect of weathering on the colour can be seen in the contrast between the topknot and the body in the freshly shorn animal on the right.

MC1R (E)

Mc1r is the second of the two alpaca base colour genes. There are two alleles hypothesised to exist at *Mc1r*. In order of dominance these are:

- E = normal colour
- e = prevents the expression of black (eumelanin)

Note that E is completely dominant over e.

There are 3 possible combinations (genotypes) of these two alleles

- EE = normal expression of Agouti
- Ee = normal expression of Agouti
- ee = no black is expressed

Mc1r is "epistatic" over *Asip*, that is, if EE or Ee is present, then *Mc1r* has no effect on *ASIP*, and it is expressed as determined by the *ASIP* genotype. If ee is present, then only the pheomelanic (non-black) portion of the *ASIP* genotype is expressed because a functional receptor is required for the Asip alleles to have an effect. Therefore, for ee animals, their base colour is the *ASIP* colour with the black removed. All ee animals have "self coloured" skin, which can range from pink to very dark brown, depending on the intensity of the pheomelanin present. Black animals with ee (i.e. aa ee) will not necessarily be white, because aa E- also produce pheomelanin that can't be seen through the intense black pigment. Therefore, an aa ee could be white to fawn in colour. An AA ee would be pink skinned white (Figure 3.7), and is the most desirable genotype for breeders wishing to produce white fibre. An A^bA^b ee would be dark fawn (aka "chestnut" as used for horses; see Figure 3.7). An a^ta^t ee would be darker fawn than an aa ee because the a^t allele has pheomelanin expressed visibly. Any "ee" animal would have slightly lighter coloured fibre than its equivalent E-, because in the agouti phenotypes, both pheomelanin and eumelanin is present in the fibre that is considered to be "brown". Therefore, the removal of the eumelanic pigment will leave only pheomelanin. Often ee animals are described as "bright" in relation to colour by alpaca breeders.

Figure 3.7: Effect of homozygous e on agouti phenotypes. Top row, E- agouti phenotypes AA, AA^b, A^bA^b, bottom row ee of the same agouti phenotype.

Figure 3.8: An A^bA^b ee alpaca clearly showing the lack of black on muzzle, around eyes, ears and lower legs. The skin is "self" coloured, but may be very dark as shown in this animal.

Classic Grey (M)

Classic Grey is a pattern, not a colour, and is probably controlled by a single, incompletely dominant gene. It acts on both pheomelanin and eumelanin, on both agouti phenotypes and ee phenotypes. There are two alleles hypothesised to exist in alpacas at this gene. In order of dominance these are:

- M = grey
- m = no effect

Note that because M is incompletely dominant to m, MM looks different to Mm, which looks different to mm.

There are 3 possible combinations (genotypes) of these two alleles

- MM = embryonic lethal
- Mm = grey
- mm = normal base colour

Classic Grey is a pattern that is overlaid on the base colour rather than a colour (the base colour is determined by *Agouti* and *Mc1r*). The symbol M is used for this gene because of its strong visual similarities to the merle gene in dogs. Based on analysis of breeding records completed by Elizabeth Paul (*personal communication*) it is thought that having two copies of classic grey (i.e. MM) is lethal because rather than the ¼: ½: ¼ white: classic grey: solid ratio of breeding results you expect from an autosomal incomplete dominant gene, classic grey x classic grey matings produce a ⅔: ⅓ classic grey: solid ratio. Animals with one copy of M have a slightly lightened version of their base colour (e.g. black = silvergrey; brown = rosegrey) on the body and they have more intense lightening on the face, neck and legs (Figure 3.9).

This paler colour is the result of a mixture of light and dark fibres, and is caused by some fibres in the affected area being paler than the base colour (i.e. the colour is diluted). The colour gets darker with age (e.g. Figure 2.10). In some cases, the darkening can be so much that the pattern is obscured. In these cases, cria photos are very useful to show the original colour (and hence the genotype) of the animal (e.g. Figures 3.10 & 3.11). Some Mm animals will have irregular shaped and sized undiluted patches of base colour in random locations on their body (Figure 3.9). If the base colour is A^bA^b or a^ta^t or heterozygotes of these, then the undiluted patches can be black and brown on the same animal. It is not clear whether the patches are part of the classic grey phenotype, or if they are caused by a separate locus. Animals with a pale base colour, i.e. white or fawn, may carry the M allele, but not be obviously classic grey.

Figure 3.9: Classic grey phenotypes. Silvergrey in the left image, and in the right image, silvergrey at the back and rosegrey at the front. Note the patch of undiluted pigment on the rosegrey, and the characteristic pattern of more intense lightening on the neck and legs compared to the body of all animals.

Figure 3.10: Cria and adult photos of the same animal. Source: Elizabeth Paul

Figure 3.11: This alpaca has been proven by breeding results to be classic grey. Source: Elizabeth Paul.

Roan (Rn)

Roan is the second most common pattern seen during the course of this project. Like Classic Grey, it also acts on both pheomelanin and eumelanin. There are two alleles hypothesised to exist in alpacas:

- rn = normal colour
- Rn = causes mixture of white and base colour fibres

Rn is dominant to rn, that is, one copy of the roan allele is enough to cause a change in the phenotype of the animal.

There are 3 possible combinations (genotypes) of these 2 alleles

- RnRn = mixed fibres
- Rnrn = mixed fibres
- rnrn = normal base colour

The tell-tale characteristics of a roan are that the body has the most mixed fibres, while the head, neck and legs remain darker (Figure 3.12). The body gets lighter with age (the opposite of classic grey). The extent of mixed fibre varies. A white animal (e.g. AA) may have the genotype Rnrn and not show the pattern. In some species, e.g. horses, RnRn is thought to be homozygous lethal (Marklund *et al.* 1999). There is not enough evidence to know if this is the case in alpacas. There is a similar lack of evidence to determine if there is any difference in phenotype between E- and ee roans.

Figure 3.12: A roan alpaca on a black base colour.

Tuxedo & Piebald

Tuxedo and piebald are two patterns that cause un-pigmented (i.e. white) regions on the animal, and are commonly referred to as spots, or white spotting patterns. They are commonly treated as two separate genes, but it is not clear if they are two separate genes, or two alleles of the same gene in alpacas. White spotting genes in horses (e.g. tobiano, sabino and dominant white) are all caused by different mutations in the KIT gene (Haase *et al.* 2007).

For both patterns, the presence of white is dominant over absence of white. Tuxedo markings tend not to occur on the body, that is, they are on the face, neck, and feet/legs (Figure 3.13). They generally do

not surround the neck. Piebald markings occur on the body as well, and are generally larger than tuxedo markings (Figure 3.14). An individual can be both tuxedo and piebald.

Figure 3.13: Alpacas with tuxedo markings. The markings on the black cria do not extend all the way around the neck. Note that the tuxedo can occur on any base colour.

Figure 3.14: A piebald marked alpaca. Note that the white on the neck extends all the way around the neck.

Appaloosa

There is insufficient data to propose an inheritance model for appaloosa in alpacas. Many coloured spot phenotypes analogous to appaloosa exist in mammals. In horses appaloosa is caused by at least two genes acting together. In pigs it is caused by an allele of *Mc1r*. Therefore, while an appaloosa-like phenotype probably exists in alpacas, there may also be other coloured spot phenotypes that are under the control of other genes, e.g. single random coloured spots that appear over time in an otherwise white animal, or unevenly distributed spots (Figure 3.16).

Figure 3.16: Coloured spots of uncertain origin

Blue Eyed White

Evidence suggests that blue-eyed white (BEW) is not a single gene trait. Breeding records collected by Elizabeth Paul suggest that BEW is caused by a combination of grey and tuxedo (or piebald). BEW can occur on an E- (black skinned) or an ee (pink skinned) base, and any agouti alleles can be present. Many BEW have coloured spots, which are usually very small (Figure 3.16). Some BEW do not have completely blue eyes, or are born with pale eyes that darken with age.

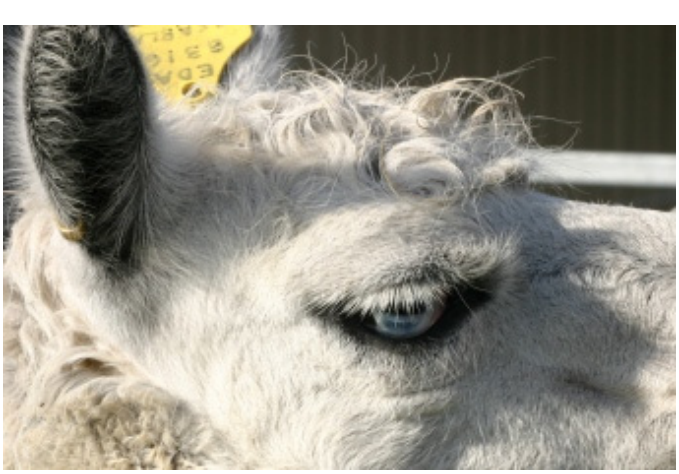

Figure 3.16: BEW with and without coloured spots.

White

White in alpacas is not a simple one-gene model. The following genotypes are hypothesised to result in white fibre:

- AA ee
- AA E-
- Aa^t or Aa ee
- Aa^t or Aa E-
- Extreme Piebald or Extreme Tux/ piebald
- BEW
- Minimal spot appaloosa

In addition, some pale bays may appear white if a classic grey allele is present.

Some of these genotypes, when mated together, will produce white (i.e. AA ee x Aa ee) and others will not. White to white matings that produce black can be caused by Aa E- x Aa E- parents, where ¼ of the offspring will be aa E-. Fawn can come from AA^b ee x AA E- where ¼ will be both E- and AA^b.

On a white or pale fawn background (i.e. AA ee, AA E-, Aa ee, or Aa E-) you can't see the effects of:

- Classic Grey
- Roan
- Piebald or Tuxedo
- Appaloosa

Therefore, many unexpected breeding outcomes can be explained by taking into account the possibility that a white parent may have additional hidden alleles present. A BEW can have any of the *Asip* alleles present, and so when mated to a dark-eyed white, could produce black, black and tan, with or without tuxedo markings, or classic grey. Alternatively an animal may appear to be white but have the genotype AA^b Mm E- and therefore produce a solid coloured offspring with full intensity fawn. If two whites carrying roan (e.g. Aa E- Rnrn) are mated, then the offspring could be white, white roan, black or black roan.

From our records classic grey is the biggest cause of confusion with colour outcomes in alpacas. Classic grey on a very pale background will produce a washed-out fawn that appears white at birth, but may develop slight colour as it ages. On a fawn, it may make the fawn appear to be light fawn. In species where gestation is short, or there is more than one offspring per pregnancy, the logical course of action for a suspect classic grey would be to mate it to a black or bay, and see if any classic grey offspring were produced (i.e. a test mating). However, this is not always practical with alpacas; it could take 4 or 5 years to confirm if a classic grey allele is present in a suspect female (Although test mating would be feasible for a male).

The four patterns listed above are all independent, that is, they can all occur in the same animal. For example, we have seen an animal that is probably both roan and classic grey on a black background, based on its parents' colours and its overall phenotype.

Physical Analysis of Colour

The spectrophotometric determination of melanin content was completed on 146 fibre samples (55 AAFL and 91 Show). There is very little difference in the amount of melanin that is produced between white, light fawn and fawn animals (Figure 3.17). Analysis of the A500/mg data (Students t-test) confirms that white is not significantly different from light fawn ($p = 0.21$) but was significantly different from fawn ($p = 0.03$) and that light fawn was not significantly different from fawn ($p = 0.15$). This data supports the hypothesis that most white alpacas are extremely diluted fawns.

The other colours had increased amounts of melanin in line with the increase in visual intensity of colour. For example, dark brown animals contained more pigment than light brown animals (Figure 3.17). There was also a significant difference observed between the show and AAFL samples in the black ($p = 1.9 \times 10^{-05}$), black brown ($p = 0.00098$) and rosegrey samples ($p = 0.0008$). These results indicated that in all cases the fibre classed by AAFL as a certain colour contained more pigment than the show animals designated to be the same colour. The most striking example of this was in the black animals, where the minimum melanin content from AAFL was higher than the maximum melanin content of the show samples. Breeders recognise that two "shades" of black exist, warm- and blue-black, with the latter being the darker of the two. There is also considerable environmental effect on black fibre, so a single animal might give a different result depending on the time of year the sample was collected.

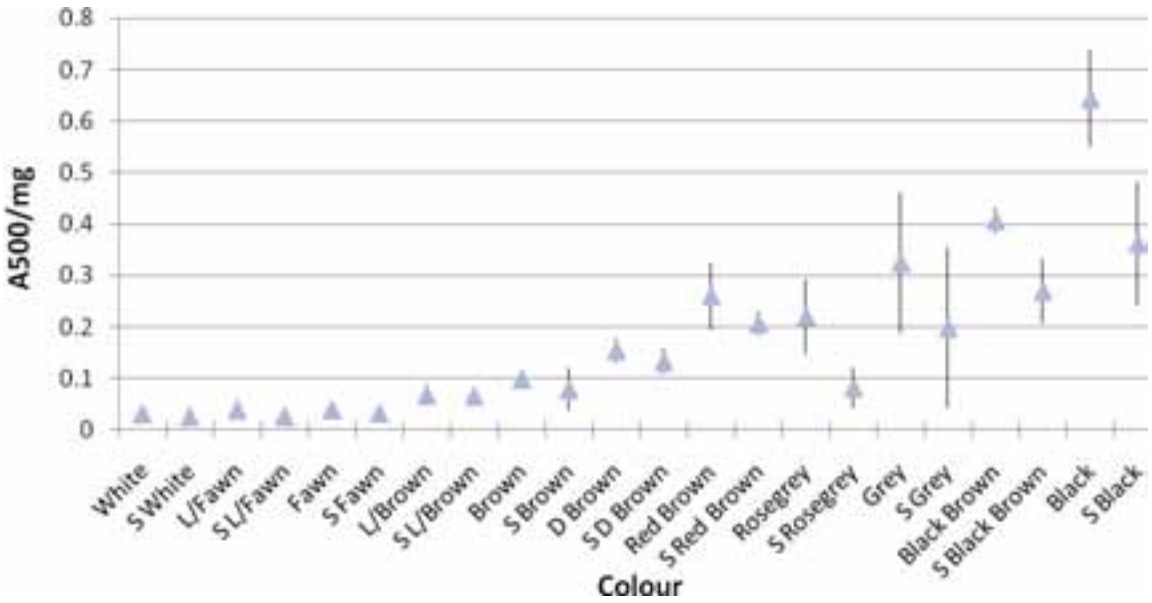

Figure 3.17: The relative amount of total melanin present in different coloured alpaca fibres. Data are shown as average (closed triangle), with maximum and minimum (shown as a vertical line). Samples are from AAFL or from show animals (preceded by an "S"). Each data point represents at least five animals.

Analysis of the 650/500 ratio, to determine the type of melanin present, again showed little difference between whites and fawns (Figure 3.18). Whites were not significantly different from light fawn ($p = 0.24$) or fawn ($p = 0.43$), nor were light fawn and fawn significantly different from each other ($p = 0.20$). Interestingly, when white was compared to all AAFL brown colours no significant difference was observed between white and brown ($p = 0.21$) and red brown ($p = 0.23$). This indicated that the major difference between the whites and these browns was the amount of pigment in the fibre, as opposed to a change in the type of pigment being produced. We found no evidence of eumelanic brown, supporting Pauls (2006) hypothesis that eumelanic brown (i.e. true genetic brown) does not exist in alpacas.

The largest range in total melanin and 650/500 ratio was seen in the grey and rosegrey samples. This is due to the variation in the description of grey. For AAFL, any dark sample with non-dark fibres was considered either rosegrey or grey. Therefore, samples with big differences in the amount of white contamination were all classed as one colour. For the rosegrey samples, the underlying *Asip* allele can significantly alter the intensity of the sample, and the amount of eumelanin versus pheomelanin. For example, an A^bA^b ee Mm rosegrey would have no eumelanin, while an A^bA^b E- Mm would have eumelanin in its fibre as well as pheomelanin.

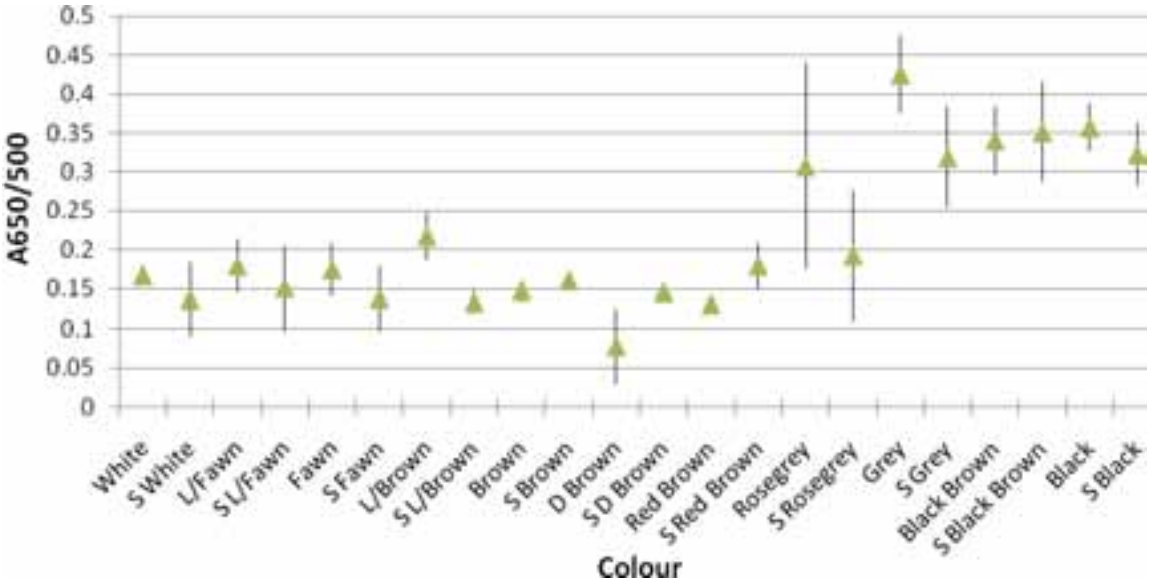

Figure 3.18: The relative proportion of type of melanin present in alpaca fibre of different colours. Data are shown as average (closed triangle), with maximum and minimum shown as a vertical line. Samples are from AAFL or from show animals (preceded by an "S"). Each data point represents at least five animals. Ratios >0.25 are indictive of eumelanin, <0.15 of pheomelanin, and intermediate (0.15 -0.25) of mixed eumealnin and pheomelanin.

If discrete colour boundaries exist in alpacas, then a plot of total melanin versus melanin type ratio should reveal separate clusters of animals, each cluster representing a colour phenotype. This was not the case (Figure 3.19). The data clearly shows that there is no difference in melanin content or type-ratio between white, light fawn, and fawn alpacas, with light brown alpacas differing only in total melanin content. These data support the hypothesis that white, light fawn and fawn alpacas are all variants of the same underlying genetic makeup. We are unable to determine if there is a difference in amount and ratio of melanins between pink-skinned and dark-skinned whites, because the AAFL samples were from unsighted animals.

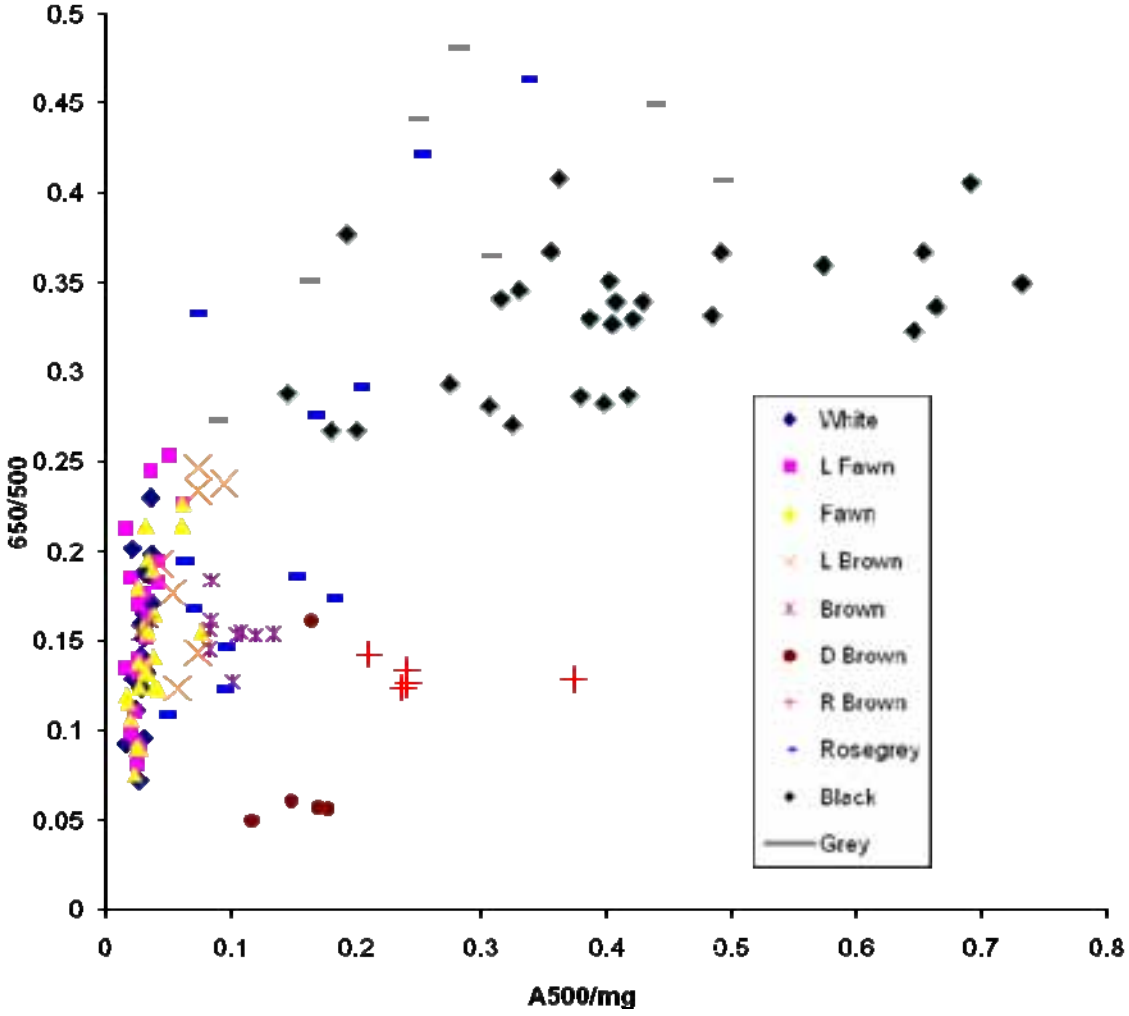

Figure 3.19: **Comparison of total melanin content with type of melanin in alpaca fibres of different colour. Data includes all AAAFL and Show samples. Total amount of melanin is on the X-axis and type of melanin is on the y-axis.**

Molecular Genetics Analysis of Colour

The Melanocortin-1 Receptor (*Mc1r*)

The melanocortin-1 receptor (*Mc1r*) gene in alpacas is a single exon gene spanning 954 base pairs (bp) and encoding a protein of only 317 amino acids. The gene is highly similar to orthologous genes in other species with 88% similarity to pig and sheep, goat 87%, cow 86%, human 85%, horse 84% and mouse 81%. The alpaca Mc1r protein was also very similar to sheep Mc1r (88%), followed by cow and cat (87%), pig (85%) and horse and dog (84%).

The alpaca *Mc1r* coding region contained 21 single nucleotide polymorphisms (SNP). Twelve of these 21 polymorphisms occurred only once in the whole dataset (C72G, A85T, G112A, G140C, A265G, T296C, T299C, C367T, T379C, T383C, A595G, and A730G) and two occurred twice (C92T and T587C). However, seven of the SNP were common enough to be considered population variants (Table 3.1). Over half of the mutant alleles did not code for any change in the protein, but three SNP (A82G, G376A and C901T) caused missense mutations T28A, G126S and R301C, respectively (Table 3.1). The SNPs A82G and T126C were linked, that is, only two combinations existed, A82/ T126 and G82/C126. Although the accepted average SNP frequency is around 1 per kilobase of DNA(Stitziel *et al.* 2003), the much higher frequency of SNP reported here is consistent with the higher number of polymorphisms reported in the *Mc1r* of other species (Jimenez-Cervantes *et al.* 2001; Majerus & Mundy 2003; FitzGerald *et al.* 2006).

Correlations between colour variants and SNP in *Mc1R* were sought. As was also found by Powell *et al.* (2009), there was no correlation between fibre colour and any SNP genotype. However, if the whole phenotype of the animal was considered (i.e. skin and fibre colour) both A82G and C901T correlated with absence of black pigment. All animals that had at least one copy of the gene version that contained A82 and C901 expressed eumelanin, while those that did not, only expressed pheomelanin. We therefore propose that the genotype A82/ C901 is the wildtype *Mc1R* allele denoted "E" and any other combination is the equivalent of "e". Where breeding records were available they supported this hypothesis. For example, the three black animals that were heterozygous *Ee* each had one black and one pheomelanic parent.

The presence of animals with pheomelanic fibre but different genotypes at *Mc1r* can be explained by considering the role of *Mc1r* in the melanocyte. If Mc1r is non-functional, then either α-MSH cannot bind or *Mc1*r cannot activate the signal cascade inside the cell even when α-MSH is bound. Inactive *Mc1r* therefore mimics the situation where ASIP is bound to Mc1r, and so only pheomelanin will be produced (Newton *et al.* 2000; Logan *et al.* 2003; Hoekstra 2006).

The change from an arginine to a cysteine driven by the C901T SNP occurs in the c-terminus of the protein. The C-terminus in Mc1r is involved in interactions with the ligand-receptor complex and correct placement of *Mc1r* within the membrane (Schoneberg *et al.* 2004; Garcia-Borron *et al.* 2005; Sanchez-Mas *et al.* 2005). In other species, polymorphisms in the C-terminus lead to significantly reduced receptor function (Everts *et al.* 2000; Garcia-Borron *et al.* 2005; Sanchez-Mas *et al.* 2005). In alpacas, the C-terminus mutation causes a change from a polar, positively charged amino acid (arginine) to a polar neutral amino acid (cysteine; Stoker 2001) a change that is likely to alter receptor interaction capacity.

Table 3.1: The colour phenotype and *Mc1r* genotypes of the alpaca samples examined in this study. "*E*" denotes the proposed wild type allele and "*e*" denotes the proposed recessive alleles at *Mc1r*. SNP in bold are the correlated with colour phenotype.

SNP Genotype									
82 (T28A)	126 (D42D)	354 (N118N)	376 (G126S)	618 (L206L)	**901 (R301C)**	933 (E311E)	**Fibre Colour**	**Eumelanin Present***	***Mc1r* alleles**
A	T	T	G	G	**C**	G	fawn	Y	*EE*
A	T	T	A	G	**C**	G	white	Y	*EE*
A	T	T	G	G	**C**	G	fawn	Y	*EE*
A	T	T	G	G	**C**	G	dk fawn	Y	*EE*
A	T	T	A	G	**C**	G	black	Y	*EE*
A	T	T	R	G	**C**	G	black	Y	*EE*
A	T	T	A	G	**C**	G	silvergrey	Y	*EE*
A	T	T	A	G	**C**	G	dk brown	Y	*EE*
A	T	T	A	G	**C**	G	med grey	Y	*EE*
A	T	T	A	G	**C**	G	black	Y	*EE*
A	T	T	A	G	**C**	G	black	Y	*EE*
A	T	T	A	G	**C**	G	black	Y	*EE*
A	T	T	A	G	**C**	G	white	?	*EE*
R	Y	Y	R	R	**Y**	R	bay	Y	*Ee*
R	Y	Y	R	R	**Y**	R	dk fawn	Y	*Ee*
R	Y	Y	R	R	**Y**	R	black	Y	*Ee*
R	Y	Y	G	R	**Y**	R	white	Y	*Ee*
R	Y	T	R	R	**Y**	R	bay	Y	*Ee*
R	Y	Y	G	R	**Y**	R	black	Y	*Ee*
R	Y	Y	A	R	**Y**	R	md fawn	Y	*Ee*
R	Y	T	R	R	**C**	R	bay	Y	*Ee*
R	Y	Y	R	R	**C**	R	black	Y	*Ee*
G	C	C	G	A	**T**	A	dk fawn	?	*ee*
G	C	C	G	A	**Y**	A	dk fawn	?	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	md fawn	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	fawn	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	C	C	G	A	**T**	A	white	N	*ee*
G	Y	C	G	A	**T**	T	rosegrey	N	*ee*
G	C	Y	G	A	**Y**	A	white	N	*ee*
G	C	C	G	R	**Y**	R	white	N	*ee*

* Y: eumelanin is present in either skin or fibre or both. N: No eumelanin is present in skin or fibre. ?: eumelanin status is unable to be definitively determined visually.
Source: (Feeley & Munyard 2009)

The other polymorphism in *Mc1r* that is linked with colour variation, A82G, causes an amino acid alteration in the extracellular loop of the N-terminus of the protein (threonine to alanine). Garcia-Borron *et al.* (2005) have found that mutations in this region of the protein in other species have functional significance, but an identical mutation has not previously been reported. Threonine is a polar hydrophilic amino acid, and alanine is a non-polar, hydrophobic amino acid (Stoker 2001). However, functional studies are required to determine the exact effect of these novel SNPs on Mc1r

activity, because modelling programs are not yet sufficiently advanced to give accurate predictions from sequence data alone.

The third missense mutation (G376A) occurs in the central section of the third transmembrane fragment of the protein, leading to a change from glycine to serine. Although the change is from a non-polar to an uncharged polar amino acid we found no correlation of either variant with colour groups. In other species, mutations in this region that affect the protein usually occur right at the boundary of the extracellular loops and the transmembrane region (Robbins *et al.* 1993; Vage *et al.* 1999). This mutation was of particular interest because mutations in this region in other species often cause activating mutations, that is, the effect would have been to cause a dominant black phenotype (Robbins *et al.* 1993; Vage *et al.* 1999). However, this mutation in alpacas clearly did not cause a dominant black phenotype and was not linked to any other phenotype (Table 3.1).

The Agouti Signalling Protein (*Asip*)

The agouti signalling protein gene (*Asip*) in alpacas consists of three coding exons named exon 2, 3 and 4. Exon 1 in other species is non-coding, and may have up to four variants per species. The entire coding region in alpacas is 402bp long, with exon 2 being 160bp, exon 3 65bp and exon 4 177 bp long. Alpaca *Asip* is longer than the human, dog and mouse homologs (3, 6 and 6bp respectively) and shorter than the cat gene (6bp). Despite these minor size differences, the gene is very similar to that in other species, with 89% similarity between alpaca and sheep, cows and goats, and 88% similarity with pigs. In alpacas Asip is 133 amino acids in length, and is again very similar to homologs in sheep, cows (83%) and horses (81%).

A total of 10 polymorphisms were found in the alpaca *Asip* coding region and introns, five SNPs in the non-coding regions, two coding regions SNPs (that caused no amino acid change), two missense SNPs and a deletion in exon 4 (Table 3.2). Three of these polymorphisms are predicted to lead to loss of function of the protein (all in exon 4; Feeley *et al.* 2011). The C-terminal 40 amino acids of Asip are coded for by exon 4, and this region is responsible for receptor binding and activity of the protein (Dinulescu & Cone 2000; Miltenberger *et al.* 2002). The 57bp deletion completely destroys a highly conserved series of disulphide bonds that stabilise the protein in its active state (McNulty *et al.* 2005; Yu & Millhauser 2007). When Asip is non-functional it cannot bind to Mc1r and therefore the cell is induced to produce only eumelanin (Figure 1.5b). We therefore propose that each of these mutations is an "a" allele at *Asip* (Table 3.3).

We used four prediction programs to model the effect of the two missense mutations on Asip function. All predicted that Asip function would be disrupted or destabilised (Feeley *et al.* 2011). The R989C mutation is also analogous to the R96C mutation that has been shown to be a loss of function mutation resulting in a recessive black phenotype in dogs (Kerns *et al.* 2004). None of the three loss-of-function alleles correlated perfectly with the black phenotype in this sample group. However, 47 of the 52 black animals had genotypes that could be considered to be homozygous loss-of-function. That is, they contained two of one mutation, or one of each of two mutations. Different mutations in other species (e.g. dogs; *Tyrp1*) combine to have the same effect on colour as two of the same mutation, and we believe that this is the case for alpacas as well (Fisher's exact test for c.325_381del57 $p = 1.3 \times 10^{-6}$; for c.292C>T $p = 0.039$; for c.353G>A $p = 0.024$; and for all combined $p = 1.92 \times 10^{-6}$).

There are four possible reasons why we found that 45% of animals carrying two "a" alleles did not exhibit a black phenotype:

- *Mc1r* is also non-functional, thus masking the effect of the *Asip* alleles (Furumura *et al.* 1996; Hoekstra 2006),

- α-Msh is non-functional, leading to no ligand signal and reversion of the cell to pheomelanin production,

- The black is being diluted by as-yet-unidentified genes (e.g. champagne in horses; Cook *et al.* 2008, chocolate in dogs, cows and sheep; Schmutz *et al.* 2002, Berryere *et al.* 2003, Gratten *et al.* 2007), and

- Inaccurate phenotype description due to ambigous terminology or environmental effects (or both).

Our *Mc1r* results show that the first option does occur (Table 3.4), but that it does not explain all cases of non-black aa alpacas.

Table 3.2: Polymorphisms identified in the alpaca ASIP gene

Polymorphism	Location	Amino Acid Effect
c. 102G>A	Exon 2	Synonymous
c. 291C>A	Exon 4	Synonymous
Exon 3 +34C>T	Intron 2	N/A
Exon 3 +56A>C	Intron 2	N/A
Exon 4 -41C>A	Intron 2	N/A
c. 292C>T	Exon 4	Arginine98Cysteine
c. 353G>A	Exon 4	Arginine118Histidine
c. 325_381del57	Exon 4	C109-R127del
Exon 4 +10C>T	3' UTR	N/A
Exon 4 +38A>G	3' UTR	N/A

Source: Feeley *et al.* 2011.

We examined the coding regions of *Asip*, but have not studied the non-coding regions in detail. In other species, the coding regions of *Asip* are involved in colour variation. Dinulescu & Cone (2000) and Fontanesi *et al.* (2010) demonstrated that *Asip* regulatory regions were responsible for cycle and location specific expression of Asip (e.g. agouti banded hairs). Three other studies (Royo *et al.* 2008, Norris and Whan 2008 and Gratten *et al.* 2009) have found that *Asip* regulatory mutations control Asip expression in sheep. We therefore expect that a close examination of *Asip* regulatory regions will uncover further mutations that affect colour in alpacas, particularly the black and tan and bay phenotypes.

Alpaca breeders can make use of these three newly discovered loss-of-function alleles, a^1, a^2 and a^3, to test their animals for the ability to produce black, and therefore reduce the incidence of non-white animals in their herds. However, breeders must be careful to maintain genetic diversity in the Australian herd, and should not cull animals based solely on this criterion. If only one parent carries and "a" allele, then the offspring cannot be black, and otherwise quality animals will be able to be retained in white herds. Conversely, if breeders wish to produce black, then they can increase the genetic diversity available to them by using non-black animals that carry a black allele. Two key points should be noted; i) that these three alleles are not the only determinants of black in alpacas, despite accounting for over 90% of all black in this study, and ii) other genes can mask the presence of black (e.g. ee) so an animal that is homozygous for a may not be black.

Table 3.3: ASIP genotypes of the three significant Exon 4 polymorphisms examined in this study.

Putative ASIP Genotype	Exon 4 Genotype			Colour	Number of Animals
	c.292C>T	c.353G>A	c.325_381del57		
	R98C	R118H	C109_R127del		
a^1a^1	CC	-	Yes	Black	21
				Brown	3
				Fawn	1
a^2a^2	TT	GG	No	Black	4
a^3a^3	CC	AA	No	Black	5
				Brown	2
				White	1
AA	CC	GG	No	Fawn	1
				White	5
a^2a^3	CT	GA	No	Black	3
				Brown	1
				Fawn	1
				White	2
Aa^3	CC	GA	No	Fawn	5
				White	3
Aa^2	CT	GG	No	Black	1
				Brown	1
				Fawn	1
				White	1
Aa^1	CC	G	Het	Black	4
				Silvergrey	1
				Brown	2
				Fawn	2
				White	1
a^1a^3	CC	A	Het	Black	2
				Brown	2
a^1a^2	CT	G	Het	Black	13
				Black & Tan	1
				Brown	2
				Fawn	1
				White	1

Source: Feeley *et al.* 2011.

Table 3.4: *Mc1r* genotypes of non-black animals with aa genotypes

Putative ASIP Genotype	Colour	Number of Animals	*Mc1r* Genotype
a^1a^1	Brown	3	EE, Ee, ee
	Fawn	1	ee
a^3a^3	Brown	2	Ee ee
	White	1	ee
a^2a^3	Brown	1	Ee
	Fawn	1	Ee
	White	2	Ee Ee
a^1a^3	Brown	2	Ee ee
a^1a^2	Black & Tan	1	Ee
	Brown	2	Ee ee
	Fawn	1	ee
	White	1	EE

Source: Feeley *et al.* 2011.

Tyrosinase Related Protein 1 (*Tyrp1*)

Alpaca *Tyrp1* is an approximately 18Kb gene consisting of 7 coding exons (exons 2 to 8) and one non-coding exon (exon 1) which is in the 5' untranslated region of the gene. The total coding region is 1617 bp long, (Exon 2 = 385bp, 3 = 323bp, 4 = 205bp, 5 = 168bp, 6 = 180bp, 7 = 147bp, and 8 = 206bp). This is the first report of the DNA sequence of alpaca *Tyrp1* exon 5, which is missing from the alpaca genome available at Ensembl. When compared to the same gene in other species, the closest relationship is between pigs (79%), cows (78%), and dolphins (78%). However, many species (including humans, chimpanzee, armadillo, horse, marmoset, panda and rabbit) have between 76 and 77% similarity, indicating that the gene is highly conserved across broad evolutionary groups. The protein encoded by this gene is 539 amino acids in length. *Tyrp1* has two copper binding sites, two cysteine rich domains, a transmembrane domain, a signal sequence and six glycosylation sites (Figure 3.20).

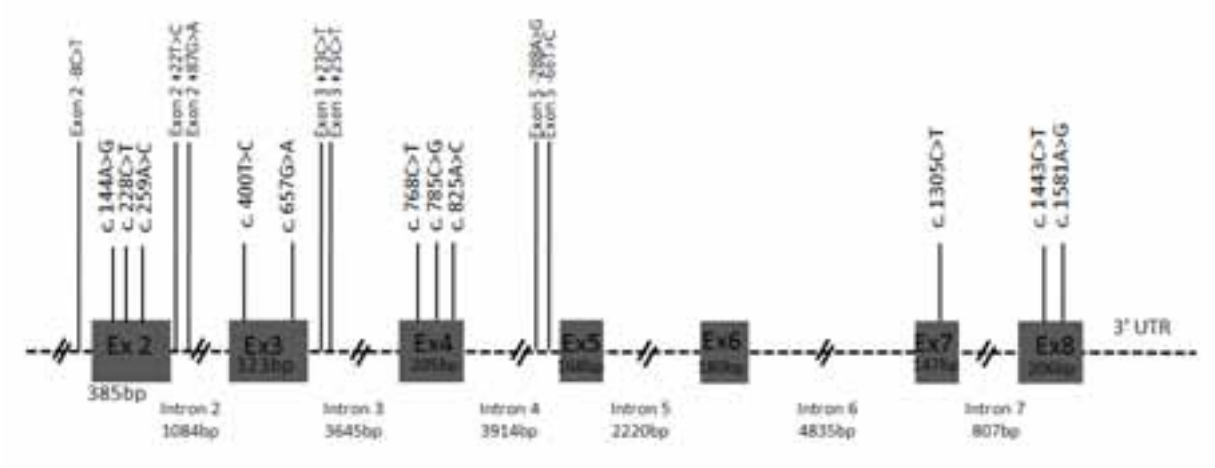

Figure 3.20: The location of newly identified alpaca *Tyrp1* mutations.

We sequenced the entire coding region of *Tyrp1* in 28 alpacas and identified 11 SNPs in the coding region (Table 3.5). A further six SNPs were found in the introns immediately flanking the exons. No mutations were found in exons 5 or 6; or in introns 5, 6 or 7 (Table 3.5). Only one of the coding exon SNPs (Exon 4:785) caused an amino acid change; from serine (S) to arginine (R). These two amino acids have different properties with serine being polar and uncharged compared to arginine which is positively charged (Blossum Matrix: -2). However, it is unlikely that this mutation has any functional effect because it falls into a region of the protein that is not highly conserved, and the region does not contain any of the functional features of the protein. The fact that all three colour groups genotyped have the mutation also suggests that the mutation does not affect colour. It is commonly assumed that SNPs that do not cause a change in amino acid can have no effect on expression. However, it has recently been shown that alternative codon usage can alter gene expression by altering miRNA binding sites (Hurst 2011). Therefore we examined all SNPs found in the coding and non-coding portion of the gene. Our sample group was selected to contain animals most likely to be eumelanic brown based on either pedigree (two black parents) or phenotype (cool chocolate brown, which is characteristic of mutations in *Tyrp1* in other species). However, none of the mutations (nor combinations of mutations) we identified were correlated with colour in our sample group.

These sequence results, along with the results of melanin analysis (Figures 3.17, 3.18, 3.19) support the hypothesis that eumelanic brown does not exist in alpacas. The existence of eumelanic brown was proposed because of breeding results where two black parents produced a brown cria. Apart from eumelanic brown, two other hypotheses can explain the production of brown cria from black parents. The first is that each of the black parents is Ee at *Mc1r*. When mated, there is a 1 in 4 chance that a cria will receive an "e" allele from each parent, becomes ee genotype at *Mc1r*, and is unable to express eumelanin. The second hypothesis is that both dominant and recessive black occur in alpacas. At a minimum, a heterozygous dominant black would have a 50:50 chance of producing a non-black cria when mated to an *Asip* aa black. In light of the evidence presented herein, it is probable that most cases of non-black cria from two black parents represent two Ee black animals in which the cria is ee, and thus non-black.

Table 3.5: Mutations in the *Tyrp1* gene in alpacas. Exons and position (base 1 is the first base of exon 1) at which SNPs were found. The single letter amino acid code for each codon is given beneath the position.

		Exon 2			Exon 3		Exon 4			Exon 7	Exon 8	
		144	228	259	400	657	768	785	825	1305	1443	1581
GROUP	PHENOTYPE	L/L	S/S	R/R	L/L	A/A	N/N	S/R	T/T	N/N	I/I	E/E
WHITE	White	AG	CT	A	T	GA	CT	C	AC	CT	CT	A
	White	AG	CT	A	T	GA	CT	CG	AC	CT	CT	A
	White	G	CT	AC	T	GA	CT	C	AC	CT	CT	A
	White	G	T	A	T	A	C	G	A	T	T	A
	White	AG	CT	A	T	GA	CT	C	AC	CT	CT	A
BLACK	Black	AG	C	AC	T	GA	CT	C	AC	..	TC	A
	Black	G	C	C	TC	G	CT	C	C	CT	C	A
	Black	G	C	C	TC	G	CT	C	C	..	C	AG
	Black	G	CT	AC	TC	GA	C	C	AC	T	C	A
BROWN	Dk Brown	G	C	C	T	G	T	C	C	C	CT	A
	Dk Brown	AG	CT	A	T	GA	CT	C	AC	CT	CT	A
	Dk Brown	G	CT	AC	T	GA	CT	C	AC	CT	CT	A
	Mid Brown	G	CT	AC	T	GA	CT	C	AC	CT	CT	A
	Brown	G	CT	AC	T	G	CT	C	AC	CT	C	A
	Dk Brown	AG	CT	A	T	G	CT	CG	AC	CT	C	A
	Warm Brown	G	CT	AC	T	G	CT	CG	AC	CT	C	A
	Red Brown	AG	CT	A	T	G	CT	C	AC	CT	C	A
	Red Brown	AG	CT	A	..	..	CT	C	AC	CT	C	A
	Dk Brown	AG	CT	A	T	G	CT	C	AC	..	C	A
	Warm Brown	A	C	A	TC	GA	CT	C	C	..	C	A
	Dk Brown	G	C	AC	T	G	CT	C	AC	CT	C	A
	Dk Brown	G	CT	AC	T	G	CT	C	..	CT	C	A
	Dk Brown	AG	C	AC	T	G	T	C	C	C	C	A
	Brown	A	C	A	T	G	T	C	..	C	C	A
	Dk Brown	AG	C	AC	TC	G	CT	C	C	CT	C	A
	Dk Brown	AG	CT	A	T	G	T	C	C	C	C	A
	Brown	A	C	A	T	G	T	C	C	C	C	A
	Brown	AG	CT	A	T	GA	CT	C	AC	CT	CT	A

Membrane Associated Transport Protein (*Matp*) & Tyrosinase (*Tyr*)

Results of the *Mc1r* analyses showed that animals with the same genotype could have different intensity of melanin, for example, from white through fawn to red-brown (aka chestnut in this report). It seemed likely, therefore, that another gene (or genes) was acting to dilute the amount of melanin present in some animals.

The palomino and buckskin phenotypes in horses occur when a chestnut or bay horse carries one defective copy of the membrane associated transport protein gene (*Matp*; Mariat *et al.* 2003*).* When two copies of the defective allele are present, a chestnut is diluted to a creamy white colour called cremello and a buckskin is diluted to a similar colour called perlino. For alpacas with the same genotype at *Mc1r*, the difference in colour was due to the differing amounts of melanin present, a situation analogous to that of chestnut, palomino and cremello in horses. Other species also have mutations in the *Matp* gene that cause dilution of colour, including the underwhite dilution in mice (Bennet & Lamoreux 2003, Newton *et al.* 2001) and Oculocutaneous albinism type 4 in humans (Newton *et al.* 2001) *Matp* was therefore a very good candidate gene to examine in alpacas for a possible dilution effect.

Tyrosinase (Tyr) is the rate limiting enzyme in the melanogenesis process (See Figure 1.3), and various mutations in *Tyr* cause albino phenotypes in cats (Imes *et al.* 2006), cattle (Schmutz *et al.* 2004), ferrets (Blaszczyk *et al.* 2007), minks (Anistoroaei *et al.* 2008), rabbits (Aigner *et al.* 2000), and rats (Blaszczyk *et al.* 2005). Many other *Tyr* alleles have been identified as causing white or near-white phenotypes, for example the chinchilla, himalayan, chinchilla-mottled and platinum phenotypes in mice (Beermann *et al.* 2004, Jackson *et al.* 1994). Therefore, we also examined *Tyr* as a possible cause of colour dilution in alpacas.

Matp has seven exons, and *Tyr* has five exons. We determined the complete sequence of all *Matp* and *Tyr* exons from 24 different alpacas. We found only two SNPs in *Matp*, one in exon 3 and one in exon 7. Only the exon 7 SNP caused an amino acid change (Table 3.6). Neither of the SNPs was correlated with colour dilution, or indeed, any colour grouping. Of the 10 new SNPs found in *Tyr*, most were in exon 1, and the rest in exons 2 and 5 (Table 3.7). Interestingly none of the exon 1 SNP caused an amino acid change, only three of the 10 SNPs did so (Table 3.7). None of the 10 Tyr mutations, nor any combination of mutations, correlated with any colour group.

Table 3.6: Single Nucleotide Polymorphisms found in the coding region of *Matp*

Exon	Nucleotide Position	Nucleotide Change	Amino Acid Change
Exon 3	843	G/A	NA
Exon 7	1526	C/T	Thr/Met

Source: Cransberg & Munyard (2011)

Table 3.7: Single Nucleotide Polymorphisms found in the coding region of *Tyr*

Exon	Nucleotide Position	Nucleotide Change	Amino Acid Change
Exon 1	78	C/A	NA
Exon 1	126	C/T	NA
Exon 1	162	C/T	NA
Exon 1	480	C/T	NA
Exon 1	513	C/T	NA
Exon 1	784	C/T	NA
Exon 2	851	G/T	Ala/Ser
Exon 5	1372	C/T	NA
Exon 5	1490	G/T	Arg/Leu
Exon 5	1498	T/C	Cys/Arg

Source: Cransberg & Munyard (2011)

The predicted Matp protein for alpacas is 526 amino acids in length and is, as expected, highly similar to other mammals, including cattle (86.6%) and humans (80.8%). Alpaca predicted Tyr is of a similar length to Matp, 530 amino acids long. Alpaca Tyr, probably because of its fundamental role in the body, is even more similar to the same protein in other species (89.4% with cattle and 88.1% with human).

Our *Mc1r* and physical analysis results suggest that colour dilution is one factor influencing variation in alpaca pigmentation. In light of the fact that neither *Matp* nor *Tyr* has any coding region polymorphisms that are correlated with colour variation in alpacas, it is logical to turn next to the regulatory regions of these genes. It is known that colour variation can be caused by changes in the promoter region of both these genes (Camp *et al.* 2003, Abul-Hassan *et al.* 2000, Graf *et al.* 2007). Regulatory region changes, that affect the extent of expression, may have a more subtle effect on colour than all-or-nothing coding region changes.

Differences in Gene Expression

Colour can be affected by alterations in the coding sequence of a protein (e.g. *Mc1r* EE vs. ee and *ASIP* AA vs. aa), but it can also be affected by the amount of that protein being expressed. Alterations in gene expression can be caused by sequence changes in the regulatory region of the gene, or by alterations in the proteins, DNA or RNA molecules that bind to the regulatory regions of the gene to control its expression levels. Each cell in an individual contains the same DNA sequence, but different cells and tissues express those genes at different times, in different amounts, or not at all. In the last few years it has become possible to obtain the sequence of almost every RNA molecule present in a transcriptome via a technique called RNA-seq. In order to discover if any expression changes were causing colour variation, we examined the RNA-seq profiles of skin cells in three groups of alpacas, pink-skinned white, bay (aka brown) and black.

The quality of the sequence reads was very high (98.9%, 98.8% and 98.6%; Table 3.8). However, only 22.5, 22.7 and 21.9% of the white, bay and black reads (respectively) were able to be mapped to the alpaca genome available on Ensembl. Similarly, only 5.9, 6.7 and 6.0% were mapped to known genes in the alpaca genome for white, bay and black animals, respectively. In contrast, the same reads mapped to the Bovine genome (also on Ensembl) matched known genes 41.9, 40.9 and 36.9% of the time. These results indicated that the alpaca genome assembly is both low-coverage, and poorly annotated. In comparison to the bovine genome, very little has been spent on the alpaca genome, and less than US$100 000 has been awarded in grants to annotate the alpaca genome. It is expected that the coverage and annotation of the alpaca genome will improve significantly in the near future due to a new sequencing project that is taking place in the US.

Table 3.8: Alpaca skin RNA-seq Read Output and Mapping Results

Criteria	White	Bay	Black
Reads obtained	9.4 million	8.9 million	7.4 million
Passed quality filter	9.3 million	8.8 million	7.3 million
Mapped to alpaca genome	2.1 million	2.0 million	1.6 million
Mapped to alpaca genes	551 000	591 000	440 000
Mapped to cow genome	5.1 million	4.4 million	2.2 million
Mapped to cow genes	3.9 million	3.6 million	2.7 million

Over 5000 alpaca genes were expressed at levels greater than 10 copies in one or more of the three skin sample sets. This high number of expressed genes reflects the heterogeneous nature of the cells making up the skin. That is, a skin sample contains epidermis, dermis, melanocytes, keratinocytes, and small numbers of adipocytes and muscle cells, all of which have different genes being expressed. Forty one known colour genes were expressed in at least one of the colours, with 25 being expressed equally in all colours (*Myo5a*, *Rab27a*, *Rab27b*, *Myc*, *Pax3*, *Slc36a1*, *Ostm1*, *Mitf*, *Atrn*, *Gnaq*, *Oca2*, *Hsp1*, *Hsp3*, *Hsp4*, *Hsp5*, *AtrnL1*, *Apc*, *Fgfr2*, *Pldn*, *Edn3*, *Slc7a1*, *Vps33a*, *Gpr161*, *Lyst* and *Cno*). Eight genes; *Rab38*, *Slc24a5*, *Tyrp1*, *Silv*, *Matp*, *Krt4*, *Oca2*, and *Tyr*, were expressed in a common pattern: highest in black, moderate in bay and lowest in white samples. This represents the first evidence that alpaca pigmentation is analogous to mouse pigmentation processes. In mice, pheomelanosomes (melanosomes filled with pheomelanin) do not have internal scaffold structures, because they lack Silv protein. They similarly lack the protein products of *Oca2* and *Matp*, the catalytic proteins Tyrp1 and Dct are almost absent, and Tyr activity is reduced (Kobayashi *et al.* 1994, Lamoreux *et al.* 1995). The amount of expression for each of these genes in alpacas is shown in Table 3.9. It is clear that alpacas also, have no expression of these genes in white animals. White alpacas are hypothesised to be ee AA, and therefore express only small amounts of pheomelanin, and these data support that hypothesis. The intermediate expression levels of these genes in bay alpacas can be explained when the aetiology of the bay colour and the location of samples from these animals are

considered. Skin samples were collected from bay animals at two places, one in the clearly brown area and one in a clearly black area. These samples were combined prior to RNA extraction and processing. Therefore, the intermediate expression levels reflect the fact that Bay samples are a mixture of pheomelanic and eumelanic skin. *Silv* is the most striking example of this differential expression, with white having only 0.5% of the expression level of black, and bay expressing only 45% the amount expressed in black.

Of these eight genes, *Rab38* is the only one with any significant expression in white samples. *Rab38* is a member of the RAS superfamily, which regulates membrane traffic by controlling vesicle formation, vesicle movement along actin and tubulin molecules, and membrane fusion. Rab38 protein has the specific action of transporting Tyrp1 protein from its place of production to the maturing melanosome. Mice that are deficient in Rab38 are chocolate in colour as a consequence of the lack of Tyrp1 in the melanosomes (Loftus *et al.* 2002). The presence of *Rab38* expression in white alpaca skin suggests that either Tyrp1 does not need to be present for Rab38 to be available for transport, or that Rab38 has another function in the cell. The difference in the level of expression between white and black alpaca skin suggests that either there is upregulation of Rab38 in black skin, or that there are fewer melanocytes in white skin than black, the latter being the most likely (c.f. similar pattern in *Ednrb and Kitlg*, also melanocyte membrane receptors; Table 3.9).

Table 3.9: Known colour genes expressed differentially in different coloured alpaca skin

Gene ID	White	Bay	Black
Rab38	86	162	201
Slc24a5	2	13	43
Tyrp1	0	14	403
Silv	2	189	423
Matp	0	17	26
Tyr	0	5	15
Oca2	0	1	4
Tyrp2	5	3	23
Krt4	62	69	82
miRNA 22	60	0	0
Kit	10	24	19
Kitlg	20	23	34
Ednrb	9	27	29
Mart-1	0	89	56
β-defensin	47	86	32

Notable absences from the list of expressed genes are *Mc1r* and *Asip*. The presence of enzymes that are expressed only in melanocytes (e.g. *Tyr*) means that melanocytes are present in alpaca skin, and the presence of colour in bay and black animals means that *Mc1r* and *Asip* must be expressed in those melanocytes. The only possible explanation for the lack of data for these genes is that they were in the subset of data that was not successfully mapped to either the alpaca or cow genomes. We consequently assembled all reads into contigs (contiguous sequences) that represented expressed genes. These contigs were then searched (via BLAST) using known *Mc1r* and *Asip* coding DNA sequences.

Significant similarities were found between *Asip* and contigs in only the white and bay skin samples; which is consistent with the hypothesis that black in alpacas is caused by non-functional Asip. Non-

functional RNA is very quickly degraded in the cell, and therefore would not be able to be detected via RNA-seq. Surprisingly, no significant matches were found between the *Mc1r* coding sequence and any contig in the sample set.

When the alpaca and cow genomes were searched specifically for *Asip* and *Mc1r* the latter was not found in either genome. In alpacas this was because there was a gap in the sequence covering over half of the *Mc1r* gene, but in the cow genome only an *Mc1r* pseudo-gene (that was almost twice the size of the real gene) was found. Significant amounts of cattle sequence remain to be mapped to the genome assembly, so the logical explanation for the lack of identified *Mc1r* in the cattle genome is that it is not yet incorporated into the assembly. However, none of these *in silico* considerations would affect the expression of *Mc1r* in alpaca skin, so it remains unclear why *Mc1r* cannot be detected via RNA-seq. One explanation for the lack of *Mc1r* expression is that there is only one Mc1r protein per melanocyte, and that because melanocytes are a small proportion of the sample, expression of *Mc1r* was not detected. However, RNA-seq is an extremely powerful tool, with much greater sensitivity than microarray or Taqman assay. Wang *et al.* (2009) present a comparison of currently used transcriptomics tools, and RNA-seq has >8000-fold greater dynamic range of sensitivity to detect expression than microarrays. The lack of ability to detect *Mc1r* in these samples therefore must be a function of technical limitations associated with the analytical software (discussed below).

Given that *Asip* sequences were present in the RNA-seq dataset, and that both the alpaca and cattle genomes have *Asip* annotated, the lack of recognition of the gene must be due to limitations in the software. Since the advent of massively high-throughput sequencing, over 20 programs have been written to assemble and analyse the data (e.g. SOAP2, Velvet, ABySS, WT, Erange), each of which has strengths and weaknesses, and new more powerful and accurate programs are constantly being developed. As the annotation and coverage of the genome, and the accuracy and thoroughness of the software improves, it is expected that even more informative data will be able to be extracted from this dataset.

Of all the genes examined in this study, three stand out as being potential causes of colour variation in alpacas: *mir22*, β-defensin and *Mart-1*. *Mir22* is not normally known as a colour-associated gene. However, it regulates *Myc* (Xiong *et al.* 2010), which is a known colour gene in mice (Lamoreux *et al.* 2010). There is a striking pattern of expression of *mir22*, where is it only expressed in white skin (Table 3.9). Micro-RNAs (miRNA) such as *mir22*, are part of the RNA silencing suite of gene regulators, and act in the cell to control expression of endogenous genes (rather than controlling exogenous genes as does siRNA). The action of miRNAs is complex, in that they act on multiple targets, each of which can, in turn, act on multiple targets. There was no difference in *Myc* expression between white, black or bay skin samples, therefore the direct action of *mir22* on *Myc* cannot be a cause of colour variation. However, the stark difference in expression of this RNA suggests that it is having some effect. *Mir22* regulates the PTEN pathway, and deficiency of PTEN is required for greying (Inoue-Narita *et al.* 2008). If *mir22* is upregulated in white skin, it may lead to accelerated "greying" or absence of pigment.

In direct contrast to the pattern seen with *mir22*, *Mart-1* expression is absent in white skin and high in bay and black skin (Table 3.9). It is tempting to speculate that *mir22* expression is negatively correlated with *Mart-1* expression, but no previous link between the two has been reported. Loftus *et al.* (2009) showed that *Mart-1* is regulated by *Mitf*. Mutation in the binding site for *Mitf* could lead to non-expression of *Mart-1*, and incorrect trafficking of *Oa1* and *Silv* to the melanosomes.

As well as known colour genes, over 60 other genes were differentially expressed in one or two of the skin samples. Many of these are well known genes (e.g. programmed cell death, creatine kinase), but there is a large number that have not been extensively studied (e.g. ring finger protein 141). It is not clear how any of the differentially expressed genes might be affecting colour. For example, small nucleolar RNAs like Snord12 are involved with post transcriptional processing of RNA, which can significantly affect the function of the targeted RNA. Examination of the functions of these differentially expressed genes is continuing.

One of the advantages of RNA-seq over TaqMan assays (real-time PCR) for determining expression levels is the ability to detect novel splice variants. Genes that appear to be equally expressed might still differ if the identity of the expressed RNA differed (i.e. have different splice variants). Unfortunately, due to the sparse annotation of the alpaca genome it was impossible to obtain information about splice variants. This aspect of the data will be revisited in the future when the available tools are better.

Implications

This research clearly shows that alpaca colour genetics is highly complex, different to sheep, mice and horses, but ultimately still explainable (same genes different mutations).

We presented a model of Mendelian inheritance that can explain almost all colour variation in alpacas. Some of the hypothesised alleles were subsequently substantiated by molecular analysis, while others remain hypothetical. A putative wild type *Mc1r* allele was found, as well as alternative alleles that render the protein non-functional, and lead to exclusive expression of pheomelanin. We identified three recessive mutations in *Asip* that render the protein non-functional, and expose breeders of white animals to the chance of producing black cria. However, these three mutations didn't explain all of the variation in this species that is hypothesised to be due to *Asip*. Further examination of the regulatory regions of *Asip* will almost certainly reveal more mutations linked with colour variation.

We found no evidence that genetic brown exists in alpacas. It is therefore inaccurate to describe them as brown, and a more appropriate name should be used e.g. bay, as in horses to describe the colour and pattern, or red, which reflects the genetics of the base colour.

The data from melanin testing revealed that most alpacas called white still have a measurable amount of melanin in their fibre, and are essentially a very dilute fawn. This further supports the hypothesis that there are multiple forms of white in the species, and that true white can be achieved by selective breeding. Skin histology of the various proposed genotypes of white alpacas in order to quantify melanocyte and melanin density and distribution, will assist in determining the underlying genetic cause of all forms of white in alpacas.

DNA tests are now available for mutations that prevent black pigment from being produced (*Mc1r*) and that show if a non-black animal carries black (*Asip*). Use of these tests will allow breeders to predict colour outcomes with more certainty. It is anticipated that more tests will become available once the underlying genetic cause of other colours (e.g. A^b and a^t) is discovered. It is also now possible to chemically test the fibre of alpacas to determine the types and amounts of melanins present. This knowledge can be used to select for "whiter" whites, distinguish between chestnut and bay, and select for "blacker" blacks.

Recommendations

It is imperative that a common, genetically based nomenclature system be introduced into the alpaca industry. Without a common language for colour, the current mis-identification and confusion will continue. Current colour descriptions for registration are not detailed enough to assist in determining inheritance patterns because they are based on only the colour of the adult fibre and don't take into account that there is probably at least two genetically distinct ways to produce a fawn, and potentially as many as five ways to get white. This has led to incorrect classification in the herd books. There is no way to determine what proportion of the herd book data is correct at a genetic level. Therefore, herd book records must be used with caution when attempting to predict colour outcomes. In addition, the colour of all cria should be recorded (including pattern) preferably by photography, so that any age-related changes can be accounted for in breeding programs, especially when animals change ownership.

Overall the results of this study indicate that in order to reduce the complexity of the colour genetics of alpacas, breeders should breed like colours together, and should not breed classic grey animals to pale solid animals. Any white animal that produces multi, classic grey or roan when mated to a solid dark animal must carry the multi, classic grey or roan gene (perhaps more than one), and should be bred with that knowledge uppermost in mind. Any animals producing non-self coloured cria can be used in breeding programs aimed at their own colour and/or for the colour(s) they produce. More research is needed on the patterns grey, roan and multi so that the extra complexity induced by these patterns can be understood completely.

The chemical melanin test is the most objective way to measure pigment in fibre, and should be used to distinguish between true white and light fawn, and blue-black and warm-black animals. Breeders wishing to produce absolutely white fibre (or very dark black) should use this test as one of the selection indices for a breeding animal. Melanin analyses should be conducted to determine if pink-skinned whites have different amount and/or ratio of melanins compared with dark-skinned whites, and whether BEW have whiter fibre that normal-eyed whites.

Despite the identification of mutations relating to base colour, there is still no genetic explanation for the variation in intensity of those colours within a given genotype. The desired intensity can be selected for in a planned breeding program; however this will take many years to have a widespread effect. In addition, it is not clear if the variation is due to genetics, environment or a combination of both. Further phenotypic and molecular research is needed to elucidate the cause of the variation.

References

ABUL-HASSAN, K., WALMSLEY, R., TOMBRAN-TINK, J. & BOULTON, M. 2000. Regulation of Tyrosinase Expression and Activity in Cultured Human Retinal Pigment Epithelial Cells. *Pigment Cell Research,* 13**,** 436-441.

AIGNER, B., BESENFELDER, U., MÜLLER, M. & BREM, G. 2000. Tyrosinase gene variants in different rabbit strains. *Mammalian Genome,* 11**,** 700-702.

ANISTOROAEI, R., FREDHOLM, M., CHRISTENSEN, K. & LEEB, T. 2008. Albinism in the American mink (*Neovison vison*) is associated with a *tyrosinase* nonsense mutation. *Animal Genetics,* 39**,** 645-648.

BEERMANN, F., ORLOW, S. J. & LAMOREUX, M. L. 2004. The *Tyr* (*albino*) locus of the laboratory mouse. *Mammalian Genome,* 15**,** 749-758.

BENNETT, D. C. & LAMOREUX, M. L. 2003. The Color Loci of Mice - A Genetic Century. *Pigment Cell Research,* 16**,** 333-344.

BERRYERE, T. G., SCHMUTZ, S. M., SCHIMPF, R. J., COWAN, C. M. & POTTER, J. 2003. TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow? *Animal Genetics,* 34**,** 169-75.

BLASZCZYK, W. M., ARNING, L., HOFFMANN, K.-P. & EPPLEN, J. T. 2005. A Tyrosinase missense mutation causes albinism in the Wistar rat. *Pigment Cell Research,* 18**,** 144-145.

BLASZCZYK, W. M., DISTLER, C., DEKOMIEN, G., ARNING, L., HOFFMANN, K. P. & EPPLEN, J. T. 2007. Identification of a tyrosinase (TYR) exon 4 deletion in albino ferrets (*Mustela putorius furo*). *Animal Genetics,* 38**,** 421-423.

CAMP, E., BADHWAR, P., MANN, G. J. & LARDELLI, M. 2003. Expression Analysis of a Tyrosinase Promoter Sequence in Zebrafish. *Pigment Cell Research,* 16**,** 117-126.

CANDILLE, S. I., KAELIN, C. B., CATTANACH, B. M., YU, B., THOMPSON, D. A., NIX, M. A., KERNS, J. A., SCHMUTZ, S. M., MILLHAUSER, G. L. & BARSH, G. S. 2007. A β-Defensin Mutation Causes Black Coat Color in Domestic Dogs. *Science,* 318**,** 1418-1423.

CHANDRAMOHAN, B., LA MANNA, V., RENIERI, C. & LA TERZA, A. 2010. Asip and Mc1r genes in alpaca. *World Congress on Genetics Applied to Livestock Production.*

COOK, D., BROOKS, S., BELLONE, R. & BAILEY, E. 2008. Missense Mutation in Exon 2 of SLC36A1 Responsible for Champagne Dilution in Horses. *PLoS Genet,* 4**,** e1000195.

CRANSBERG, R. & MUNYARD, K. A. 2011. Polymorphisms detected in the tyrosinase and matp (slc45a2) genes did not explain coat colour dilution in a sample of Alpaca (*Vicugna pacos*). *Small Ruminant Research,* 95**,** 92-96.

DINULESCU, D. M. & CONE, R. D. 2000. Agouti and Agouti-related Protein: Analogies and Contrasts. *Journal of Biological Chemistry,* 275**,** 6695-6698.

EVERTS, R. E., ROTHUIZEN, J. & VAN OOST, B. A. 2000. Identification of a premature stop codon in the melanocyte-stimulating hormone receptor gene (MC1R) in Labrador and Golden retrievers with yellow coat colour. *Animal Genetics,* 31**,** 194-9.

FEELEY, N. L., BOTTOMLEY, S. & MUNYARD, K. A. 2011. Three novel mutations in ASIP associated with black fibre in alpacas (*Vicugna pacos*). *The Journal of Agricultural Science,* FirstView**,** 1-10.

FEELEY, N. L. & MUNYARD, K. A. 2009. Characterisation of the melanocortin-1 receptor gene in alpaca and identification of possible markers associated with phenotypic variations in colour. *Animal Production Science,* 49**,** 675-681.

FITZGERALD, L. M., FRYER, J. L., DWYER, T. & HUMPHREY, S. M. 2006. Effect of MELANOTAN(R), [Nle4, D-Phe7]-α-MSH, on melanin synthesis in humans with MC1R variant alleles. *Peptides,* 27**,** 388-394.

FONTANESI, L., TAZZOLI, M., RUSSO, V. & BEEVER, J. 2010. Genetic heterogeneity at the bovine *KIT* gene in cattle breeds carrying different putative alleles at the *spotting* locus. *Animal Genetics,* 41**,** 295-303.

FRANK, E. N., HICK, M. V. H., GAUNA, C. D., LAMAS, H. E., RENIERI, C. & ANTONINI, M. 2006. Phenotypic and genetic description of fibre traits in South American domestic camelids (llamas and alpacas). *Small Ruminant Research,* 61**,** 113-129.

FURUMURA, M., SAKAI, C., ABDEL-MALEK, Z., BARSH, G. S. & HEARING, V. J. 1996. The Interaction of Agouti Signal Protein and Melanocyte Stimulating Hormone to Regulate Melanin Formation in Mammals. *Pigment Cell Research,* 9**,** 191-203.

GARCIA-BORRON, J. C., SANCHEZ-LAORDEN, B. L. & JIMENEZ-CERVANTES, C. 2005. Melanocortin-1 receptor structure and functional regulation. *Pigment Cell Research,* 18**,** 393-410.

GRAF, J., VOISEY, J., HUGHES, I. & VAN DAAL, A. 2007. Promoter polymorphisms in the MATP (SLC45A2) gene are associated with normal human skin color variation. *Human Mutation,* 28**,** 710-717.

GRATTEN, J., BERALDI, D., LOWDER, B. V., MCRAE, A. F., VISSCHER, P. M., PEMBERTON, J. M. & SLATE, J. 2007. Compelling evidence that a single nucleotide substitution in *TYRP1* is responsible for coat-colour polymorphism in a free-living population of Soay sheep. *Proceedings: Biological Sciences,* 274**,** 619-626.

GRATTEN, J., WILSON, A. J., MCRAE, A. F., BERALDI, D., VISSCHER, P. M., PEMBERTON, J. M. & SLATE, J. 2008. A Localized Negative Genetic Correlation Constrains Microevolution of Coat Color in Wild Sheep. *Science,* 319**,** 318-320.

HAASE, B., BROOKS, S. A., SCHLUMBAUM, A., AZOR, P. J., BAILEY, E., ALAEDDINE, F., MEVISSEN, M., BURGER, D., PONCET, P.-A., RIEDER, S. & LEEB, T. 2007. Allelic Heterogeneity at the Equine *KIT* Locus in Dominant White (*W*) Horses. *PLoS Genet,* 3**,** e195.

HOEKSTRA, H. E. 2006. Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity,* 97**,** 222-234.

HURST, L. 2011 The sound of silence. *Nature,* 471, 582-583.

IMES, D. L., GEARY, L. A., GRAHN, R. A. & LYONS, L. A. 2006. Albinism in the domestic cat (*Felis catus*) is associated with a tyrosinase (TYR) mutation. *Animal Genetics,* 37**,** 175-178.

INOUE-NARITA, T., HAMADA, K., SASAKI, T., HATAKEYAMA, S., FUJITA, S., KAWAHARA, K., SASAKI, M., KISHIMOTO, H., EGUCHI, S., KOJIMA, I., BEERMANN, F., KIMURA, T., OSAWA, M., ITAMI, S., MAK, T. W., NAKANO, T., MANABE, M. & SUZUKI, A. 2008. Pten Deficiency in Melanocytes Results in Resistance to Hair Graying and Susceptibility to Carcinogen-Induced Melanomagenesis. *Cancer Research,* 68**,** 5760-5768.

JACKSON, I. J., BUDD, P., HORN, J. M., JOHNSON, R., RAYMOND, S. & STEEL, K. 1994. Genetics and Molecular Biology of Mouse Pigmentation. *Pigment Cell Research,* 7**,** 73-80.

JIMENEZ-CERVANTES, C., GERMER, S., GONZALEZ, P., SANCHEZ, J., SANCHEZ, C. O. & GARCIA-BORRON, J. C. 2001. Thr40 and Met122 are new partial loss-of-function natural mutations of the human melanocortin 1 receptor. *FEBS Letters,* 508**,** 44-48.

KOBAYASHI, T., URABE, K., ORLOW, S. J., HIGASHI, K., IMOKAWA, G., KWON, B. S., POTTERF, B. & HEARING, V. J. 1994. The Pmel 17/silver locus protein. Characterization and investigation of its melanogenic function. *J. Biol. Chem.,* 269**,** 29198-29205.

LAMOREUX, M. L., DELMAS, V., LARUE, L. & BENNETT, D. C. 2010. *The Colors of Mice: A Model Genetic Network*, Wiley-Blackwell.

LIN, J. Y. & FISHER, D. E. 2007. Melanocyte biology and skin pigmentation. *Nature,* 445**,** 843-850.

LOFTUS, S. K., BAXTER, L. L., BUAC, K., WATKINS-CHOW, D. E., LARSON, D. M. & PAVAN, W. J. 2009. Comparison of melanoblast expression patterns identifies distinct classes of genes. *Pigment Cell & Melanoma Research,* 22**,** 611-622.

LOGAN, D. W., BRYSON-RICHARDSON, R. J., PAGAN, K. E., TAYLOR, M. S., CURRIE, P. D. & JACKSON, I. J. 2003. The structure and evolution of the melanocortin and MCH receptors in fish and mammals. *Genomics,* 81**,** 184-191.

MAJERUS, M. E. N. & MUNDY, N. I. 2003. Mammalian melanism: natural selection in black and white. *Trends in Genetics,* 19**,** 585-588.

MARIAT, D., TAOURIT, S. & GUERIN, G. 2003. A mutation in the MATP gene causes the cream coat colour in the horse. *Genetics Selection Evolution,* 35**,** 119-133.

MARKLUND, S., MOLLER, M., SANDBERG, K. & ANDERSSON, L. 1999. Close association between sequence polymorphism in the *KIT* gene and the roan coat color in horses. *Mammalian Genome,* 10**,** 283-288.

MCGREGOR, B. A. 2006. Production, attributes and relative value of alpaca fleeces in southern Australia and implications for industry development. *Small Ruminant Research,* 61**,** 93-111.

MCNULTY, J. C., JACKSON, P. J., THOMPSON, D. A., CHAI, B., GANTZ, I., BARSH, G. S., DAWSON, P. E. & MILLHAUSER, G. L. 2005. Structures of the Agouti Signaling Protein. *Journal of Molecular Biology,* 346**,** 1059-1070.

MILLER, S. A., DYKES, D. D. & POLESKY, H. F. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucl. Acids Res.,* 16**,** 1215-.

MILTENBERGER, R. J., WAKAMATSU, K., ITO, S., WOYCHIK, R. P., RUSSELL, L. B. & MICHAUD, E. J. 2002. Molecular and Phenotypic Analysis of 25 Recessive, Homozygous-Viable Alleles at the Mouse agouti Locus. *Genetics,* 160**,** 659-674.

NEWTON, J. M., COHEN-BARAK, O., HAGIWARA, N., GARDNER, J. M., DAVISSON, M. T., KING, R. A. & BRILLIANT, M. H. 2001. Mutations in the Human Orthologue of the Mouse underwhite Gene (uw) Underlie a New Form of Oculocutaneous Albinism, OCA4. *The American Journal of Human Genetics,* 69**,** 981-988.

NEWTON, J. M., WILKIE, A. L., HE, L., JORDAN, S. A., METALLINOS, D. L., HOLMES, N. G., JACKSON, I. & BARSH, G. S. (2000) Melanocortin 1 receptor variation in the domestic dog. *Mammalian Genome,* 11, 24-30.

NORRIS, B. J. & WHAN, V. A. 2008. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research,* 18**,** 1282-1293.

OZEKI, H., ITO, S., WAKAMATSU, K. & THODY, A. J. 1996. Spectrophotometric Characterization of Eumelanin and Pheomelanin in Hair. *Pigment Cell Research,* 9**,** 265-270.

PAUL, E. 2006. Colour Review. *Australian Alpaca Association.* Adelaide Convention Centre.

POWELL, A. J., MOSS, M. J., TREE, L. T., ROEDER, B. L., CARLETON, C. L., CAMPBELL, E. & KOOYMAN, D. L. 2008. Characterization of the effect of Melanocortin 1 Receptor, a member of the

hair color genetic locus, in alpaca (*Lama pacos*) fleece color differentiation. *Small Ruminant Research,* 79**,** 183-187.

REES, J. L. 2003. Genetics of hair and skin color. *Annual Reviews in Genetics,* 37**,** 67-90.

ROBBINS, L. S., NADEAU, J. H., JOHNSON, K. R., KELLY, M. A., ROSELLI-REHFUSS, L., BAACK, E., MOUNTJOY, K. G. & CONE, R. D. 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell,* 72**,** 827-834.

ROYO, L. J., ALVAREZ, I., ARRANZ, J. J., FERNANDEZ, I., RODRIGUEZ, A., PEREZ-PARDAL, L. & GOYACHE, F. 2008. Differences in the expression of the ASIP gene are involved in the recessive black coat colour pattern in sheep: evidence from the rare Xalda sheep breed. *Animal Genetics,* 39**,** 290-293.

SANCHEZ-MAS, J., SANCHEZ-LAORDEN, B. L., GUILLO, L. A., JIMENEZ-CERVANTES, C. & GARCIA-BORRON, J. C. 2005. The melanocortin-1 receptor carboxyl terminal pentapeptide is essential for MC1R function and expression on the cell surface. *Peptides,* 26**,** 1848-1857.

SCHMUTZ, S., BERRYERE, T., CIOBANU, D., MILEHAM, A., SCHMIDTZ, B. & FREDHOLM, M. 2004. A form of albinism in cattle is caused by a tyrosinase frameshift mutation. *Mammalian Genome,* 15**,** 62-67.

SCHMUTZ, S. M., BERRYERE, T. G. & GOLDFINCH, A. D. 2002. TYRP1 and MC1R genotypes and their effects on coat color in dogs. *Mamm Genome,* 13**,** 380-7.

SCHONEBERG, T., SCHULZ, A., BIEBERMANN, H., HERMSDORF, T., ROMPLER, H. & SANGKUHL, K. 2004. Mutant G-protein-coupled receptors as a cause of human diseases. *Pharmacology & Therapeutics,* 104**,** 173-206.

SPONENBERG, P. 2001. Some Educated Guesses on Colour Genetics of Alpacas. *Alpaca Registry Journal.*

STITZIEL, N. O., TSENG, Y. Y., PERVOUCHINE, D., GODDEAU, D., KASIF, S. & LIANG, J. 2003. Structural Location of Disease-associated Single-nucleotide Polymorphisms. *Journal of Molecular Biology,* 327**,** 1021-1030.

STOKER, H. 2001. *General, Organic and Biological Chemistry,* Boston, Houghton Mifflin.

TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology,* 28**,** 511-515.

VAGE, D. I., KLUNGLAND, H., LU, D. & CONE, R. D. 1999. Molecular and pharmacological characterization of dominant black coat color in sheep. *Mamm Genome,* 10**,** 39-43.

WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet,* 10**,** 57-63.

XIONG J, DU Q, & LIANG, Z. 2010. Tumor-suppressive microRNA-22 inhibits the transcription of E-box-containing c-Myc target genes by silencing c-Myc binding protein. *Oncogene* **29** (35): 4980–8.

YU, B. & MILLHAUSER, G. L. 2007. Chemical disulfide mapping identifies an inhibitor cystine knot in the agouti signaling protein. *FEBS Letters,* 581**,** 5561-5565.

Inheritance of White Colour in Alpacas

— Identifying the genes involved —

by Kylie Munyard

Publication No. 11/074

This report describes the research conducted as part of the alpaca colour genetics project to identify the genes involved in the inheritance of white colour in alpacas. Three approaches were used (Mendelian, physical and genetic) in an attempt to unravel the mystery surrounding colour inheritance in alpacas.

This project has successfully identified key mutations in genes that lead to differences in fibre colour in alpacas. Other genes, which play a role in colour variation in other species, were cleared of involvement in colour variation in alpacas. Through extensive observational analysis a model for Mendelian inheritance of the major colours was developed. In combination, these findings provide breeders with information that allows them to make informed colour breeding choices.

Contact RIRDC:
Level 2
15 National Circuit
Barton ACT 2600

PO Box 4776
Kingston ACT 2604

Ph: 02 6271 4100
Fax: 02 6271 4199
Email: rirdc@rirdc.gov.au
web: www.rirdc.gov.au
Bookshop: 1300 634 313