

The Pennsylvania State University

The Graduate School

College of the Liberal Arts

HUMAN SKIN PIGMENTATION VARIATION:

A PHENOTYPIC, GENOTYPIC, AND EVOLUTIONARY PERSPECTIVE

A Thesis in

Anthropology

by

Heather Lynne Norton

© 2005 Heather Lynne Norton

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2005

The thesis of Heather Lynne Norton was reviewed and approved* by the following:

Mark D. Shriver
Associate Professor of Biological Anthropology
Thesis Advisor
Chair of Committee

Kenneth M. Weiss
Evan Pugh Professor of Biological Anthropology and Genetics

Jeffery A. Kurland
Associate Professor of Biological Anthropology and Human
Development

Kateryna Makova
Assistant Professor of Biology

Dean R. Snow
Professor of Anthropology
Head of the Department of Anthropology

*Signatures are on file in the Graduate School

ABSTRACT

Skin pigmentation varies both within and between human populations. Broad patterns of variation between human populations tend to follow geographic clines that are highly correlated with the amount of UVR striking the surface of the earth. Many have suggested that the broad patterns of human pigmentation variation reflect changes due to natural selection while others feel that sexual selection is a better explanation. The purpose of this dissertation is to improve current understanding of localized patterns of pigmentation variation in Island Melanesia, the genes underlying that variation, and the potential role that natural selection has played in determining global patterns of pigmentation.

To assess phenotypic variation across a localized geographic region I analyzed quantitative measures of skin and hair pigmentation in 1135 Island Melanesians. While skin pigmentation variation in the region has likely been constrained by natural selection the patterns of variation that I observed within and between islands suggests that localized patterns of variation may have been influenced by genetic drift and/or sexual selection. The population history of the region is characterized by intermittent waves of migration and long periods of isolation and drift in some islands. This has led to complex population substructure across the region that complicates the interpretation of genotype-phenotype associations that I observed between the *OCA2* and *ASIP* genes and skin pigmentation and the *MC1R* gene and hair pigmentation. To control for this stratification in future studies I proposed a method using a panel of neutral markers that will be able to differentiate between different subpopulations in the region

Pairwise locus-specific F_{ST} is used to detect signals of selection in six pigmentation candidate genes in six geographically diverse populations. Four of the six genes show high pairwise F_{ST} values that may be indicative of population divergence due to selection. All four were confirmed to be associated with normal skin pigmentation variation using admixture mapping. I also typed these four loci in the CEPH-diversity panel to confirm these signals of selection. Two of these genes, *OCA2* and *ASIP*, are associated with pigmentation differences between lightly and darkly pigmented populations. The distribution of the light and dark alleles at *OCA2* A355G in the CEPH-diversity panel suggests that polymorphism at *OCA2* may be quite old; the high frequency of the light *OCA2* allele in the Khoisan sample is consistent with this interpretation. The distribution of the light and dark alleles at *ASIP* A8818G in this sample set suggests that the *ASIP* dark allele is most common in Africa and occurs at much lower frequencies elsewhere, including at least some (but not all) Island Melanesian populations. *MATP* and *TYR* show strong signals of European-significant divergence, suggesting that lighter skin in European populations may be due to mutations independent of those associated with lighter pigmentation in East Asian populations. This is consistent with the genotyping results in the CEPH-diversity panel, which demonstrated that the light allele in each gene occurs at its highest frequency in Europe and declines rapidly as one moves east and south.

This dissertation has highlighted the potential for localized variation in human skin and hair pigmentation to have been shaped by genetic drift or sexual selection while still being constrained by natural selection. It has also demonstrated the importance of being able to control for population stratification in genotype-phenotype association

studies. Lastly, it used a locus-specific F_{ST} -based method to identify pigmentation candidate genes that may have been subject to selection.

TABLE OF CONTENTS

List of Figures.....	ix
List of Tables	xi
Acknowledgements.....	xiii
Chapter 1 Background and Introduction.....	1
Introduction.....	2
Phenotypic Variation	6
Pigmentation and Racial Classification.....	10
Origins of Phenotypic Variation.....	13
Photoprotection.....	14
Vitamin D Hypothesis	15
Folic Acid Hypothesis	16
Concealment.....	17
Thermoregulation	17
Disease Resistance.....	19
Melanin Biosynthesis.....	19
Pigmentation Candidate Genes.....	22
Conclusions.....	27
Literature Cited.....	28
Chapter 2 Skin and Hair Pigmentation Variation in Island Melanesia.....	46
Introduction.....	47
Materials and Methods	50
Measurement	51
Island Classification	53
Linguistic Classification.....	53
Neighborhood Classification.....	55
Latitude/longitude Classification.....	56
Statistical Analysis	58
Results.....	58
Sex-based Differences	58
Effects of Increasing Adult Age.....	59
Linguistic Phylum	60
Island	62
Neighborhood.....	63
Latitude and Longitude.....	67
Interaction Effects	68
Discussion.....	68
Sex-based Differences.....	68

Age	70
Linguistic Phylum	72
Island and Neighborhood Variation	75
Latitude and Longitude.....	85
Conclusions.....	88
Acknowledgements.....	89
Literature Cited.....	90
 Chapter 3 Pigmentation Candidate Gene Variation in Island Melanesia	 99
Introduction.....	100
Materials and Methods	101
Samples.....	101
Laboratory Methods	103
Statistical Methods	104
Results.....	106
Phenotypic Variation	106
Genotypic Variation	108
Hardy Weinberg Equilibrium.....	112
Hierarchical Locus-Specific F Statistics	113
Genotype-phenotype Associations	114
Discussion.....	117
Phenotypic Variation	117
Genotypic Variation	117
Hardy Weinberg Equilibrium.....	121
Hierarchical Locus-Specific F Statistics	121
Genotype-phenotype Associations	123
Conclusions.....	126
Literature cited.....	129
 Chapter 4 A Candidate Gene Study for Signals of Natural Selection Using Pairwise Locus-specific F_{ST}	 138
Introduction.....	139
Materials and Methods	144
Samples.....	144
Affymetrix Chip	144
Pigmentation Candidate Gene Typing	144
MATP Genotype-phenotype Association	145
Marker Selection	146
PCR and Genotyping Methods.....	146
F_{ST} Estimation and Percentile Rank Calculation.....	146
Phylogenetic Tree Construction	147
Genotype-phenotype Associations at <i>MATP</i>	147

Results.....	148
Pairwise Locus-specific F_{ST}	148
<i>MATP</i> C374G Genotype-phenotype Association.....	158
Pigmentation SNP Allele Frequencies in CEPH-Diversity Samples	159
Discussion.....	163
Conclusions.....	169
Literature Cited.....	173
 Chapter 5 Conclusions and Future Directions	 180
Introduction.....	181
Ascertainment Bias in SNP Selection.....	182
Identification of SNPs vs. Haplotypes Associated with Pigmentation Variation	184
Changes in Pigmentation Over the Course of Human Evolution	185
Future Directions	192
Pigmentation Genotype-phenotype Associations in Island Melanesia	192
Identifying Genotype-phenotype Linkage in Other Admixed Populations...	194
Confirming the Signal of Selection in <i>MATP</i> and <i>TYR</i>	195
Distinguishing Between Sexual and Natural Selection	196
Conclusions.....	198
Literature Cited.....	200

LIST OF FIGURES

Figure 1.1 : The Raper-Mason pathway of melanin biosynthesis. Note that tyrosinase is the rate-limiting enzyme in the production of both eumelanin and pheomelanin.....	20
Figure 1.2 : Melanocyte-keratinocyte complex. Pigmentation candidate genes are identified in red.....	22
Figure 2.1 : Map of the study region in Island Melanesia. Samples were obtained from the islands of New Guinea, New Britain, New Ireland, Mussau, New Hanover, and Bougainville (denoted *)......	50
Figure 2.2 : Skin M index for 27 of the 28 neighborhoods reported on in this study (the Sepik of New Guinea are not included). Intensity of the circle marking each population on the map corresponds to pigmentation as measured by the M index.....	76
Figure 2.3 : Distributions of skin M index values for the Bougainvilleans measured in this study, African Americans and African Caribbeans measured by Shriver et al. (2003)......	78
Figure 2.4 : Hair M index for 27 of the 28 neighborhoods reported on in this study (the Sepik of New Guinea are not included). Intensity of the circle marking each population on the map corresponds to pigmentation as measured by the M index.....	84
Figure 3.1 : Allele frequencies for the four SNPs polymorphic in Island Melanesia according to island, pooled across the region, and in each of the five comparison populations.....	111
Figure 4.1 : Pairwise locus-specific F_{ST} distributions for two population pairs: Island Melanesians vs. Native Americans and Europeans vs. South Asians.....	149
Figure 4.2 : Population tree based on average pairwise locus-specific F_{ST} across the 11,078 SNPs typed on the Affymetrix 10K chip.....	153
Figure 4.3 : Locus-specific tree for <i>ASIP</i> A8818G.....	154
Figure 4.4 : Locus-specific tree for <i>OCA2</i> A355G.....	155
Figure 4.5 : Locus-specific tree for <i>TYR</i> A192C.....	156
Figure 4.6 : Locus-specific tree for <i>MATP</i> C374G.....	157

- Figure **4.7**: Distribution of *ASIP* A8818G in the CEPH-diversity sample set. The “light” allele is shown in yellow..... 160
- Figure **4.8**: Distribution of *OCA2* A355G in the CEPH-diversity sample set. The “light” allele is shown in yellow..... 160
- Figure **4.9**: Distribution of *TYR* A192C in the CEPH-diversity sample set. The “light” allele is shown in yellow..... 161
- Figure **4.10**: Distribution of *MATP* C374G in the CEPH-diversity sample set. The “light” allele is shown in yellow. 161

LIST OF TABLES

Table 2.1: Sample composition of Austronesian and Papuan speakers.....	55
Table 2.2: Mean skin and hair M index values for males and females in each linguistic phylum	60
Table 2.3: ANOVA results comparing islands, neighborhoods, neighborhoods within phyla and islands	62
Table 2.4: Mean skin and hair M index values for islands ¹ and neighborhoods ²	64
Table 2.5: Mean skin and hair M index values for males and females on each island.....	65
Table 2.6: Mean skin and hair M index values for males and females in each neighborhood	66
Table 2.7: Mean M and UVMED values by island	87
Table 3.1: Sample composition of Austronesian and Papuan speakers.....	101
Table 3.2: Raw skin and hair M index values for neighborhood and island samples. Papuan-speaking neighborhoods are shown in boldface	102
Table 3.3: Marker amplification and digestion information	105
Table 3.4: Results of ANOVAs comparing skin and hair M index between neighborhoods and linguistic phyla on each island	107
Table 3.5: Mean allele frequencies according to island, neighborhood, linguistic phylum and comparison samples.....	110
Table 3.6: Significance tests for departures from Hardy Weinberg Equilibrium (deficiency of heterozygotes)	113
Table 3.7: Hierarchical F_{ST} values in the region at large, and within each island	114
Table 3.8: Genotype-phenotype associations within the sample-at-large, and on each island	116
Table 3.9: Significant secondary interaction effects on skin and hair pigmentation ..	116
Table 4.1: Possible outcomes of pairwise locus-specific F_{ST} comparisons between populations.....	143

Table 4.2: Allele frequencies by population	151
Table 4.3: Average pairwise F_{ST} for the Affymetrix 10K WGA chip and locus-specific F_{ST} for four pigmentation candidate SNPs. P-values are shown in parentheses	152

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Dr. Mark Shriver for his support, guidance, and friendship during both my graduate and undergraduate careers at Penn State. I am also grateful to my committee members: Kenneth Weiss, Jeffrey Kurland, and Kateryna Makova. All of them have generously given their time and provided me with valuable insights during this process, and I believe that the final product is better for it.

I also owe a deep debt of gratitude to Dr. Jonathan Friedlaender for allowing a young, inexperienced graduate student to tag along on a field expedition at the last possible minute. Over the years he has provided advice, discussion, friendship, and a strong kick in the pants when required. I am also grateful to his wife Françoise Friedlaender for her support and assistance with the production of the maps in Chapter Four. Dr. Anne Buchanan has been a great source of encouragement, as I know she is and has been to so many other graduate students.

I would also like to thank the many members of the Shriver lab, past and present, who have provided assistance and stimulating intellectual discussions: Dr. Carrie Pfaff, Dr. Esteban Parra, Dr. Sonia Dios, Dr. Carolina Bonilla, Jennifer Wagner, Indrani Halder, Marc Bauchet, Abigail Bigham, Xianyun Maio, and Tamiko Brown. Corey Sparks has always generously provided his time and statistical assistance whenever it was required, and for that I am extremely grateful. I also thank the people of Papua New Guinea who participated in the this research—their hospitality and interest made fieldwork a joy, as did the assistance and friendship of George Koki of the Papua New Guinea Institute of

Medical Research. This research was funded in part by a Wenner Gren Dissertation Grant (#7138) and a grant to Mark Shriver from the NIH/NHGRI (#HG002154).

I must also thank my parents, John and Maria Norton, and my siblings Matthew and Rachel, for the love and support they have always shown me—this work would not have been possible without them. Lastly, I thank my husband, Christopher Pierce, who has been with me on this journey from the beginning, and never once thought of turning back.

Chapter 1
Background and Introduction

Introduction

Following a recent common origin in Africa humans migrated through the world colonizing a diverse array of environments that included arid deserts, humid jungles, temperate zones, and arctic regions. The wide geographic range of *Homo sapiens* is especially intriguing considering that most genetic evidence suggests that anatomically modern humans (AMH) arose in Africa only ~ 150,000-200,000 years ago (Cann et al., 1987; Vigilant et al., 1991; Stoneking et al., 1992; Goldstein et al., 1995; Stoneking et al., 1997; Seielstad et al., 1999; Thomson et al., 2000; Underhill et al., 2001). Fossil evidence supports this (Singer and Wymer, 1982; Rightmire and Deacon, 1991; Clark et al., 2003; White et al., 2003; Haile-Selassie et al., 2004), although earlier hominids (e.g. *H. erectus*, *H. georgicus*, *H. neanderthalensis*) have been found outside of Africa at sites ranging from Indonesia (Dubois, 1926; Sartono, 1971), Caucasus Georgia (Gabunia and Vekua, 1995; Vekua et al., 2002), and the Middle East/Europe (Boule, 1908; Arsuaga et al., 1993; Rak et al., 1994).

The success with which *Homo sapiens* has settled every continent except for Antarctica suggests that humans possess an unusual ability to adapt to a wide variety of environmental conditions. Although surely aided by the development of human material culture, including tools and clothing, this adaptation to novel environments has likely also been the result of genetic adaptations and associated phenotypic changes that have improved fitness in particular regions. Over time this adaptation can cause populations

living under different environmental conditions to diverge from each other more quickly at loci under selection. Reproductively isolated populations will also diverge due to the effects of genetic drift. Drift will affect all regions of the genome equally, unlike natural selection which affects only those loci associated with a particular adaptation.

While the different selective regimes encountered in diverse environments can cause populations to diverge at specific loci, sometimes geographically separated populations experience similar environmental pressures. Will this similarity in environment lead to genetic convergence at specific loci or will the populations develop independent adaptations? One example of this phenomenon is the widespread distribution of malaria across parts of Africa, the Mediterranean, and the southwest Pacific. Indigenous populations in these regions have acquired a number of different mutations that confer full or partial resistance to malaria. While the diverse mechanisms through which humans have independently adapted to malarial environments are well-studied, there are likely other traits that show similar patterns of convergence.

Human skin pigmentation is a phenotypic trait that can vary dramatically across different environmental regimes, suggesting that differences in pigmentation may be the result of adaptation (Cowles, 1959; Blum, 1961; Wasserman, 1965; Loomis, 1967; Walter, 1971; Branda and Eaton, 1978; Mackintosh, 2001). Others (Darwin, 1871; Diamond, 1992) suggest that this variation may be better explained by mate choice, or sexual selection. While phenotypic differences in pigmentation between populations have been widely acknowledged, the evolutionary and genetic origins of these differences are less well-understood. It is also unknown if similarity in pigmentation phenotype

necessarily implies similarity at the level of one pigmentation gene, or if instead this phenotypic similarity is the result of independent adaptation at several genes.

This dissertation seeks to address questions such as these by exploring variation in pigmentation phenotype, identifying possible genotype-phenotype associations, and searching for signals of genetic adaptation in pigmentation candidate genes. As such, it is comprised of three main articles. The first (Chapter Two) describes skin and hair pigmentation variation in Island Melanesia in terms of the melanin (M) index, an objective measure of the amount of melanin contributing to skin or hair pigmentation. Given the generally high levels of ultraviolet radiation (UVR) in Island Melanesia, we would expect natural selection to have favored darker skin. While skin pigmentation is generally dark in the region relative to other populations there is still a remarkable amount of skin and hair pigmentation variation present. This variation is explored in light of geography, linguistics, and inferences about population history in the region. This work stresses the fact that it is difficult to assign a general pigment “type” to all Island Melanesians. It also suggests that even under strong selective pressures, genetic drift and/or sexual selection may play significant roles in shaping phenotypic variation within a region.

Chapter Three describes allele frequencies at ten different SNPs in six pigmentation candidate genes in a subset of the Island Melanesian sample described in Chapter Two. Allele frequencies are compared between groups within the region, as well as to a set of five global comparison populations that include Native Americans, Europeans, West Africans, East Asians and South Asians. In addition to reporting allele frequencies, within-region heterogeneity at pigmentation loci is assessed using

hierarchical locus-specific F -statistics. Primary and secondary associations between candidate genes and pigmentation phenotype are discussed. Population stratification in the region makes interpreting significant associations difficult, and a plan to control for this heterogeneity is proposed.

In the third article (Chapter Four) pairwise locus-specific F_{ST} is used to identify putative signals of natural selection in pigmentation candidate loci in six geographically diverse populations. A number of hypotheses have been put forward in an attempt to explain the global distribution of skin pigmentation as a function of adaptation to environments that vary in UVR levels. However, the role of natural selection in shaping variation in pigmentation genes has been rarely explored (for an exception, see Rana et al. 1999; Harding et al. 2000; and Makova et al. 2001). In an effort to address these questions I calculated pairwise locus-specific F_{ST} for ten pigmentation candidate SNPs in six geographically diverse human populations. High pairwise locus-specific F_{ST} values (relative to an empirical distribution) are indicative of population divergence that is due most likely to the effects of natural or sexual selection. Such values will be of particular interest in comparisons of populations that are different in pigmentation, as they will provide genetic evidence that population differences may be explained by selection (pending confirmation that the SNP in question is associated with functional variation in pigmentation). High pairwise locus-specific F_{ST} values will also be of great interest in comparisons of populations that are similar in pigmentation phenotype, as they may indicate possible cases of phenotypic convergence (once again, assuming that a functional effect can be demonstrated).

As an introduction to the results presented in Chapters Two, Three, and Four the remainder of this chapter will provide background on the current state of knowledge regarding human pigmentation. This includes a brief history of the study of this variation and its role in racial taxonomies, a review of the primary selection-based hypotheses of pigmentation variation, and a description of the pigmentation candidate SNPs typed in Chapters Three and Four. Chapter Five will address how future research in the variation and evolution of human skin pigmentation should proceed to answer some of the questions raised in this dissertation.

Phenotypic Variation

One of the earliest surviving depictions of human skin pigmentation variation can be found in Egyptian artwork. A statue from the Fourth Dynasty (dating to 2163 BC) shows a medium-toned Prince Rahotep with his very light-skinned consort Nefret (Kahane, 1967). Artwork on the walls of the tomb of another Egyptian pharaoh, Sethos I, (dating to 1300 BC) depicts a Libyan, Nubian, Asian, and Egyptian all with varying skin tones consistent with modern phenotypic variation between these groups (Holubar, 1996). These works suggest that by at least 4,000 years ago pigmentation showed extensive variation within and between populations. More recent evidence comes from Classical Greek artwork dating to the sixth century BC that depicts dark-skinned Africans and lightly pigmented Mediterraneans (Snowden, 1970). The Bonampak Murals from Chiapas, Mexico (a Classic Maya site dating to 250 A.D.-850 A.D.) illustrate

pigmentation variation among the Maya (Miller, 1986), although the differences are much smaller than those observed in the Egyptian or Greek artwork.

While early artistic portrayals may help us to place a lower bound on the time since pigmentation variation was first observed in different regions they do not provide us with information on early beliefs regarding the origins of this variation. However, the writings from Classical Greek and Rome provide some insights to ancient hypotheses explaining these phenotypic differences between human populations. For example, the Greek Herodotus suggested that “Africans were black with the heat” encountered in their land (Herodotus, 415 BC). The belief that pigmentation was influenced by climatic conditions was a common one that was later expanded upon by the astrologer Ptolemy, who claimed that “the demarcation of national characteristics is established in part by entire parallels and angles, through their position relative to the ecliptic and the sun” (Ptolemy, 140 AD). Ptolemy noticed that pigmentation seemed to vary in a predictable way as one traveled away from tropical regions: individuals living in hot Africa had darkly pigmented skin, while those living in colder northern climes generally had lighter coloring. Although much of Ptolemy’s *Tetrabiblos* would today fall under the category of pseudoscience, the relationship between pigmentation and climate remains the basis for many of today’s hypotheses about normal variation in skin pigmentation. The ancient Romans rightly suspected that pigmentation was a hereditary trait, and paid particular attention to the progeny of African and Greek/Roman couples. They also traced pigmentation variation in these admixed families and noted the effects of this admixture in later generations. For example, Pliny observes that a marriage between an African

man and a Greek woman resulted in a “white” child. Later, when that woman had a child, it was “as black as its grandfather” (Pliny, 77 AD).

In Judeo-Christian traditions a popular explanation for the dark skin pigmentation of Africans was that it resulted from a curse from God. The curse of Cain, in which Cain was permanently marked by God because he had deceived his father and murdered his brother, was sometimes used to explain the darker pigment of Africans. An alternate story suggests that Noah’s grandson, Canaan (son of Ham), was the subject of God’s curse, and that he and his descendents (rather than Cain’s) bore the mark of God’s displeasure. Religious explanations of pigmentation variation may have played an important role in shaping European colonial attitudes towards other populations. It is notable that it was often the darker skin of these groups that merited explanation, rather than the lighter skin of Europeans.

Rather than seeking out religious explanations for the darker skin of Africans, one 18th century physician suggested that it might be the result of a widespread infection with leprosy (Rush, 1799). While in some cases leprosy is known to produce a dark tinge to the skin, Rush took his leprosy diagnosis one step further by stating that leprosy was also associated with insensibility of the nerves, vitiligo, strong venereal desires, big lips, flat noses, and “woolly heads”, all traits that he associated with Africans (Rush, 1799). As further evidence for his argument that black skin was actually a disease, Rush claimed that a white

woman living with an African male in North Carolina began to assume characteristics of Africans.

Contemporaries of Rush focused on the role that environment might play in determining skin color. S.S. Smith, a moral theologian, noted the gradation of skin color that existed from tropical to northern latitudes and suggested that perhaps the dark skin tone of Africans was due to a combination of effects from the hot tropical sun on skin (much like a permanent suntan), as well as the transformation of bile (one of the four humors) from yellow to black when exposed to the sun and air. Although he recognized the connection between environment and pigmentation, Smith was writing prior to Darwin's *Origin of Species* and did not view adaptation to different environmental conditions as a natural process. Instead, he suggested that the human ability to adapt was a gift from God, intended to help humans colonize all regions of the earth (Smith, 1788).

Knowledge of the scientific causes of pigmentation variation in the 18th and 19th centuries was often influenced by studies of pigmentation abnormalities. Due to their easily visible effects albinism and vitiligo were often studied. In American society, it was not uncommon for depigmentation of the skin to be associated with the development of behavioral traits more commonly found only in "white" populations. In light of the fact that racial hierarchies at the time were associated as much with character traits as with physical appearance this is not necessarily surprising. For example, in 1839 Marcy described a case of "the Cape May albinos", a family in which a woman bore three

normal African children and three albino children. Marcy described the parents of this family as far more industrious and virtuous than “the majority of the common Negroes of the neighborhood” (Marcy, 1839), attributing their fine moral character at least in part to whatever hereditary factor predisposed their children to whiteness of the skin.

Not only western populations associated differences in pigmentation with differences in character. The San Blas Cuna, a Native American population indigenous to Caribbean islands off the coast of Panama, have a high incidence of albinism (11/17,000 compared to a worldwide frequency of 1/17,000; Keeler, 1953). Albinos in the San Blas populations are referred to as “Moon Children”, the name being derived from the belief that their abnormal pigmentation is the result of their mothers spending too much time gazing at the moon during pregnancy (Keeler, 1953). Interestingly, the Cuna often describe albinos in their population as “lazy”, and cite as evidence for this their preference for “basket weaving, singing, and choir directing” rather than for fishing or other forms of outdoor labor (Keeler, 1968). However, moon children are also believed to have a deeper spiritual connection to the gods, so their depigmentation is also associated with elevated social status in Cuna society.

Pigmentation and Racial Classification

Pigmentation has featured prominently in racial taxonomies and has colored modern day perceptions of race. Early racial definitions suggested that races represented discrete physical types that could be described largely on the basis of phenotypic differences. In his *Tetrabiblos* Ptolemy described three main human groups: his own

(the Greeks), the dark Ethiopians of the south, and the fair Scythians to the north. In addition to affecting skin pigment, Ptolemy also believed that the different climates that these groups inhabited affected their characters. Not surprisingly, the Greeks, living in a temperate climate, were described as more civilized than either the Ethiopians or the Scythia (Ptolemy, 140 AD).

One of the first western instances of a clear-cut “scientific” racial typology comes from Carolus Linnaeus. Linnaeus identified four human races: *Homo sapiens afer*, *Homo sapiens europeus*, *Homo sapiens asiaticus*, and *Homo sapiens americanus*. The Linnaean divisions were based primarily on geography, and each of the four groups was characterized in terms of pigmentation, character, and stance. For example, Asians were described as “sallow, avaricious, and easily distracted” while Europeans were “white, gentle, and inventive” (Linnaeus, 1735). Although some believe that the Linnaean racial characterizations were not inherently hierarchical (Gould, 1981), the scientific association of character with physical features may have lead many to believe that these were permanently linked together and that they were true characterizations of large groups of people.

Linnaeus’ student, the anthropologist Johann von Friedrich Blumenbach eventually added a fifth group (Malay) to his teacher’s racial taxonomy. Blumenbach argued that the difference between these races was the result of humans spreading out from the point of creation and adopting different habits and customs in different regions, undergoing what he described as “degeneration”. As he believed white Caucasians resided in the region closest to the point of Creation, Blumenbach felt that Caucasians represented the least degenerate of the human populations. As one moved further and

further from this Creation point, the degree of degeneration encountered in populations increased. This degeneration would lead to variation in physical traits, including skin color. For Blumenbach differences in skin pigmentation could be attributed to “the abundance of carbon in the human body, which, when it is excreted with the hydrogen through the corium, and precipitated by the contact of the atmospheric oxygen, becomes imbedded in the Malpighian mucus” (Blumenbach, 1865). Because he believed that the reactions of carbon with atmospheric oxygen could vary in different environmental regions, Blumenbach felt that pigmentation was a racial trait that could, over time, change in response to environment. Interestingly, although he identified what he believed were five main races Blumenbach believed that variation between these groups was in fact continuous: “innumerable varieties of mankind run into one another by insensible degrees” (Blumenbach, 1865). This continuous variation between populations is just one of many problems commonly cited with racial classifications of humans today.

More recently genetic data has been used to address questions of race. Lewontin demonstrated that only 6.3% of human genetic diversity could be attributed to variation between racial groups; 8.3% was due to variation within racial groups, and 85.4% was due to variation within populations (Lewontin, 1972). Others have expressed similar findings, usually using Wright’s F_{ST} —average values for this and related statistics across multiple populations and loci is about 15% (Wright, 1951; Cavalli-Sforza et al., 1994; Barbujani et al., 1997; Jorde et al., 2000; Tishkoff and Verrelli, 2003; Watkins et al., 2003). However, loci that may be involved in localized adaptation to different environmental regimes (such as those involved in skin pigmentation) may have F_{ST}

values higher than this average. Locus-specific F_{ST} values for pigmentation genes will be explored in Chapter Four.

Origins of Phenotypic Variation

Differences in pigmentation between and within human populations have long been noted. But how did this variation arise? In his *The Descent of Man and Selection in Relation to Sex* Charles Darwin suggested that the variation in human pigmentation was too extensive to be explained by natural selection alone (Darwin, 1871). Instead, he favored sexual selection, and cited numerous examples from around the world suggesting that each region had its own ideal type of beauty. This idea has more recently been revived by Jared Diamond (Diamond, 1992). Others have analyzed the vast quantity of phenotypic data in the literature, as well as reports of mate preference as described in the Human Relations Area Files in an effort to investigate this hypothesis further (Van den Berghe and Frost, 1986; Frost, 1988; Frost, 1994). These studies are suggestive of a universal preference for lighter females, and indeed many quantitative phenotypic studies have reported a (slight) quantitative difference in skin pigmentation between males and females (Barnicot, 1958; Tobias, 1961; Conway and Baker, 1972; Byard and Lees, 1982; Harvey, 1985). Although sometimes significant these differences are never very large, and it is difficult to know if they are reflecting true biological variation or are instead an artifact of the measurement site used or different male/female activity levels associated with differential UVR exposure levels (Lock-Andersen and Wulf, 1997). Interestingly, females typically experience a decrease in pigmentation levels (relative to males) at

puberty, and pigmentation may increase during pregnancy. This has caused some (Van den Berghe and Frost, 1986) to suggest that a lighter skin may be a sign of fecundity in females.

Although there has been some research into the role of sexual selection in influencing pigmentation variation (largely within populations, rather than across broad geographic regions) the vast majority of hypotheses explaining the distribution of pigmentation values across the human species invoke natural selection. Six of the more prominent of these will be reviewed below.

Photoprotection

It has long been believed, although only more recently proven, that the melanin found in darkly pigmented individuals provides some level of protection from the UVR produced by the sun (Walter, 1971; Pathak and Fitzpatrick, 1974; Roberts and Kahlon, 1976). These protective effects of melanin are thought to explain why most populations found in high UVR regions are darker than those found in higher latitudes, where UVR is lower. Protection provided by highly melanized skin would eliminate or minimize the painful effects of sunburn as well as reduce the risk of skin cancer (Robins, 1991). From a reproductive fitness point of view the weakest point in the photoprotection hypothesis is that skin cancer is typically a disease that strikes late in life, after the primary period of reproduction for most humans. This makes the strength of selection in this case rather weak (Blum, 1961). The benefits of avoiding a painful sunburn in tropical environments should not be ignored, and it is easy to imagine the difficulties that would be faced by

relatively pale-skinned hominids in a tropical savannah environment. In fact, albinos in regions where UVR is high often suffer from reduced fitness and develop skin cancers early in life (Okoro, 1975; Luande et al., 1985)—this may explain the Cuna Moon Children’s preference for “lazy” indoor activities.

Vitamin D Hypothesis

Vitamin D₃ is a necessary requirement for proper utilization of dietary calcium for bone mineralization. Deficiencies in Vitamin D₃ have been linked to an increased risk of rickets. Vitamin D₃ can be obtained through diet, traditionally in foods such as fish liver, or it can be synthesized within the body through the photoactivation of intermediate compounds in the dermis of the skin. As early humans migrated from the tropical regions to upper latitudes they would be exposed to less and less UVR. Murray (1934) theorized that a highly melanized skin would be maladaptive in such an environment, since the high levels of melanin would presumably inhibit the synthesis of the amount of Vitamin D₃ needed to prevent rickets. Under the Vitamin D hypothesis, human populations migrating to regions of lower UVR experienced positive selection for variants that reduced or altered the amount and/or type of melanin in the skin, hence reducing skin color. Over thirty years later Loomis (1967) revived Murray’s Vitamin D hypothesis and also suggested that light skin might be maladaptive in high UVR regions because it could lead to toxic over-production of Vitamin D₃. However, Holick et al. (1987) have demonstrated that regardless of UV exposure skin-mediated Vitamin D₃ synthesis cannot

exceed certain limits. As such, overproduction of Vitamin D₃ is not a selective pressure that could lead to an increase in skin pigmentation in high UVR regions. Some populations living in low UVR regions have skin pigmentation that is darker than expected given the predictions of the Vitamin D hypothesis (Jablonski and Chaplin, 2000). This suggests that either these groups are obtaining their Vitamin D₃ through dietary means (as Murray suggested for the Inuit), or that perhaps vitamin D₃ synthesis is not critical in determining skin pigmentation. Alternatively, these groups may not have lived in low UVR regions long enough for selection to bring about an adaptive depigmentation of the skin.

Folic Acid Hypothesis

The folic acid hypothesis is also based on the ability of melanin to at least partially block UVR. UVR exposure can lead to the breakdown of folic acid in the human body. A large body of literature beginning in the 1980's has strengthened the link between folic acid deficiencies and neural tube birth defects (Laurence et al., 1981; Bower and Stanley, 1989; MVSR, 1991). An association between folic acid deficiencies and increased errors of nondisjunction during spermatogenesis has also been observed (Mathur et al., 1977). The folic acid hypothesis (Branda and Eaton, 1978) suggests that there is strong selective pressure to maintain a darkly pigmented skin in higher UVR regions to avoid the breakdown of folic acid. Recently Flemming and Copp (1998) identified a link between folate deficiency and neural tube birth defects in the mouse, creating renewed interest in this hypothesis (Jablonski and Chaplin 2000). This

hypothesis predicts that selection would favor darker skin in regions of higher UVR; outside of these regions pigmentation may vary due to relaxed functional constraints.

Concealment

Cowles (1959) argued that before their mastery of tools and hunting techniques early hominids would have been easy prey for large predators. As such, he argues that a darker skin would have been advantageous for concealment in a tropical jungle environment. This hypothesis would be more convincing if we had evidence that Homo, particularly later species of Homo, actually evolved in jungle-like environments. However, paleoclimatic evidence indicates a shift towards more arid conditions in Africa starting around 2.8 million years ago (deMenocal, 1995), suggesting that it is more likely that members of the genus Homo evolved in environments more similar to gallery forest and savannah than tropical jungle. Although we cannot be certain when man lost his body hair (but see Kittler et al. 2003 and Rogers et al. 2004 for interesting attempts to answer this question) it would have to have been quite early in hominid evolution for Cowles' concealment hypothesis to be a viable option.

Thermoregulation

Gloger's rule states that animals with darker pigmentation are typically found in warm, moist climates, while lighter pigmentation is common in animals living in cold, dry regions. Since darker colors are thought to absorb more heat, this makes the

abundance of darkly pigmented organisms (including humans) in hotter regions seem paradoxical, as it may actually compromise thermoregulation. Darker coats in desert animals such as goats and birds have been reported to aid in a reduction of metabolic rate that may assist in survival through periods of food scarcity (Ohmart and Lasiewski, 1971; Dmiel et al., 1980), suggesting that the benefits of a slower metabolism far outweigh the problems presented by overheating in these environments. The relevance of this to humans is unclear. It may be that in regions of high UVR efficient sweating is a more important feature than pigmentation (Robins, 1991). Cowles (1959) suggested that Gloger's rule may be based more on the fact that hot, humid regions tend to be darker (the result of extensive jungle vegetation). As such, Gloger's rule is really better described as an adaptation for concealment purposes (see above) than thermoregulation. Comparisons of European and African populations in physical tasks under a variety of environmental conditions suggested that heat acclimatization rather than skin pigmentation was more important in determining thermoregulation capabilities (Strydom and Wyndham, 1963). However, other studies suggest that African populations may sweat less under hot and humid conditions, which could be an important factor in maintaining proper hydration levels (Baker, 1958; Ladell, 1964; Riggs and Sargent, 1964). Reported associations between darker skin and frostbite have been taken to suggest that depigmented skin may be an adaptation to colder regions, (Post et al., 1975; Post and Rao, 1977) although exactly how melanin content would modify physiological response to cold was unknown.

Disease Resistance

Wasserman (1965) argued that the darkly pigmented skin of tropical populations is the result of adaptations against tropical disease, rather than tropical climate. This hypothesis suggests that in order to survive the diseases prevalent in such environments individuals would develop adaptations that would lead to a decrease in corticosteroids, thus increasing reticuloendothelial system (RES) activity. This increase in RES activity would confer greater resistance to infections and parasites. As a side effect, the decrease in corticosteroids would promote an increase in the pituitary secretion of adrenocorticotrophic hormone (ACTH) and melanocyte-stimulating hormone (MSH), resulting in an increase in skin pigmentation. Similarly, Mackintosh (2001) has suggested that the primary function of melanin and melanocytes may be to prevent parasitic, fungal, and bacterial infections of the skin (Mackintosh, 2001), which are common in tropical regions. Recent studies have demonstrated a link between immunosuppression and UVR (Schacter et al., 1983; Weitzen and Bonavida, 1984; McMichael and Hall, 1997; Termorshuizen et al., 2002; Staples et al., 2003), and some pigmentary disorders, such as Gricelli syndrome and Chediak-Higashi disease, are also characterized by immune system deficiencies.

Melanin Biosynthesis

Melanin is the primary pigment of the skin (and hair), although other molecules such as hemoglobin also affect skin coloration and appearance. In order to understand how natural selection may have shaped variation in skin pigmentation it is essential to

understand the underlying biochemistry of melanin production and the physiology of the melanin-producing cells of the body, the melanocytes. Melanin in mammalian skin and hair is produced in two forms: brown/black eumelanin and red/yellow pheomelanin. The melanin biosynthesis pathway, also known as the Raper-Mason pathway, is illustrated in Figure 1.1. The initial steps in the production of both eumelanin and pheomelanin require the hydroxylation of tyrosine to 3, 4-dihydroxyphenylalanine (DOPA) and the oxidation of DOPA to DOPAquinone by tyrosinase. At this point, the melanin synthesis pathway diverges; pheomelanin is produced from the metabolites of 5-S-cysteinylDOPA and eumelanin is produced from the metabolites of DOPochrome.

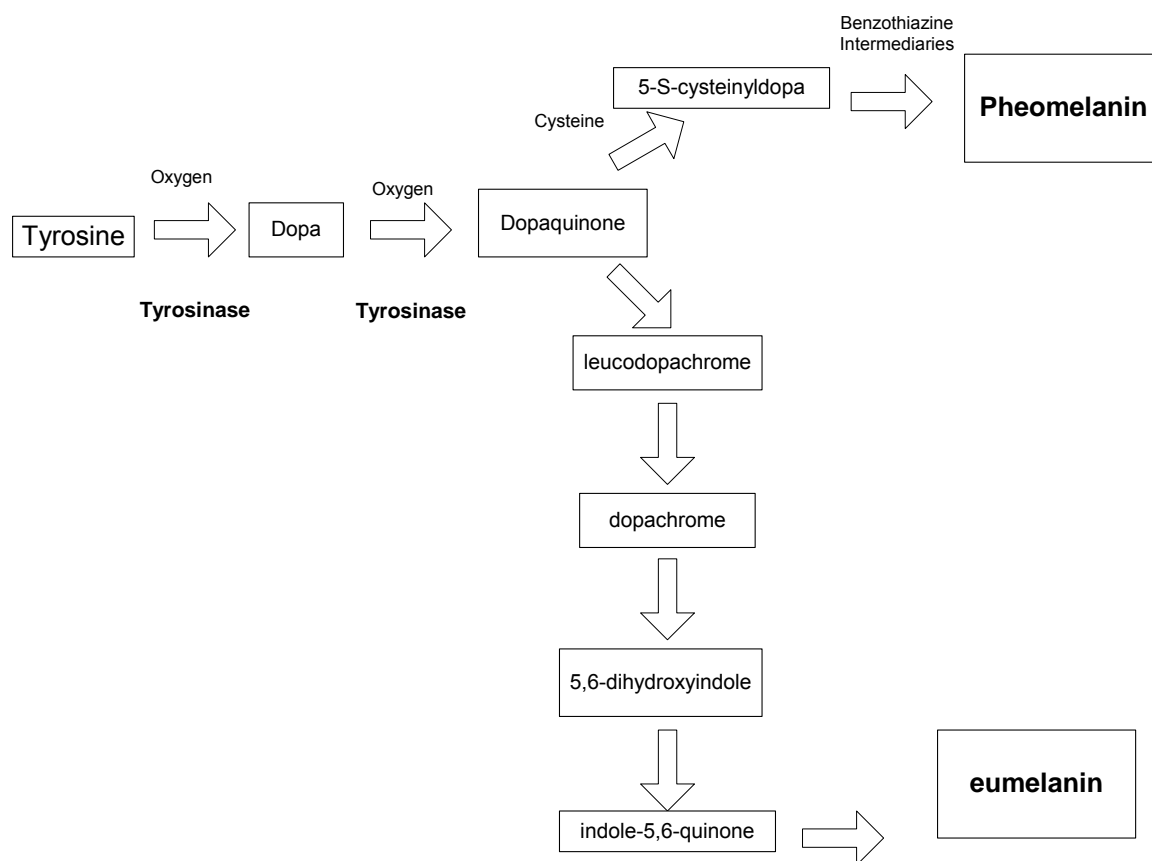


Figure 1.1: The Raper-Mason pathway of melanin biosynthesis. Note that tyrosinase is the rate-limiting enzyme in the production of both eumelanin and pheomelanin

Melanin production takes place inside specialized cells found in the basal layer of the skin known as melanocytes. After it is synthesized, melanin is packaged into membrane-bound organelles known as melanosomes. Eumelanin is stored in larger ellipsoidal melanosomes with a highly complex glycoprotein matrix and pheomelanin is stored in smaller, spherical melanosomes with a less complex glycoprotein matrix (Sturm et al., 2001). Nordlund et al. reported that individuals of “diverse ethnic background” demonstrate quantifiable differences in both melanosome distribution as well as the degree of melanization (Nordlund et al., 1998). They observed that darker skin tends to be comprised of greater numbers of the larger melanosomes, while lighter skin is associated with sparsely distributed aggregates of the smaller melanosomes. Following maturation, melanosomes are transported out of the melanocyte cell and into keratinocytes. The exact transfer mechanism is unknown, but two primary candidates are phagocytosis or endocytosis (Westbroek et al., 2001). Once they have been transmitted to the keratinocytes the melanosome membranes will degrade, allowing the melanin to diffuse throughout the cell. In many cases the melanin may form a protective “cap” over the nucleus of the keratinocyte, providing protection from UVR (Montagna and Carlisle, 1991). A diagram of the melanocyte-keratinocyte complex with pigmentation candidate genes highlighted is shown in Figure 1.2.

Exposure to UVR can stimulate melanogenesis as part of the tanning response (Friedmann and Gilchrest, 1987), most likely due to the resulting DNA damage or the DNA repair intermediates that are produced following UVR exposure (Agar and Young, 2005). Some studies have reported that in the presence of pheomelanin and its precursor, 5-S-cysteinyldopa, UVR is highly

mutagenic (Harsanyi et al., 1980; Koch and Chedekel, 1986). This suggests that the type of melanin produced may have an effect on risk of UVR-induced damage such as melanoma (Schmitz et al., 1995; Sturm, 1998).

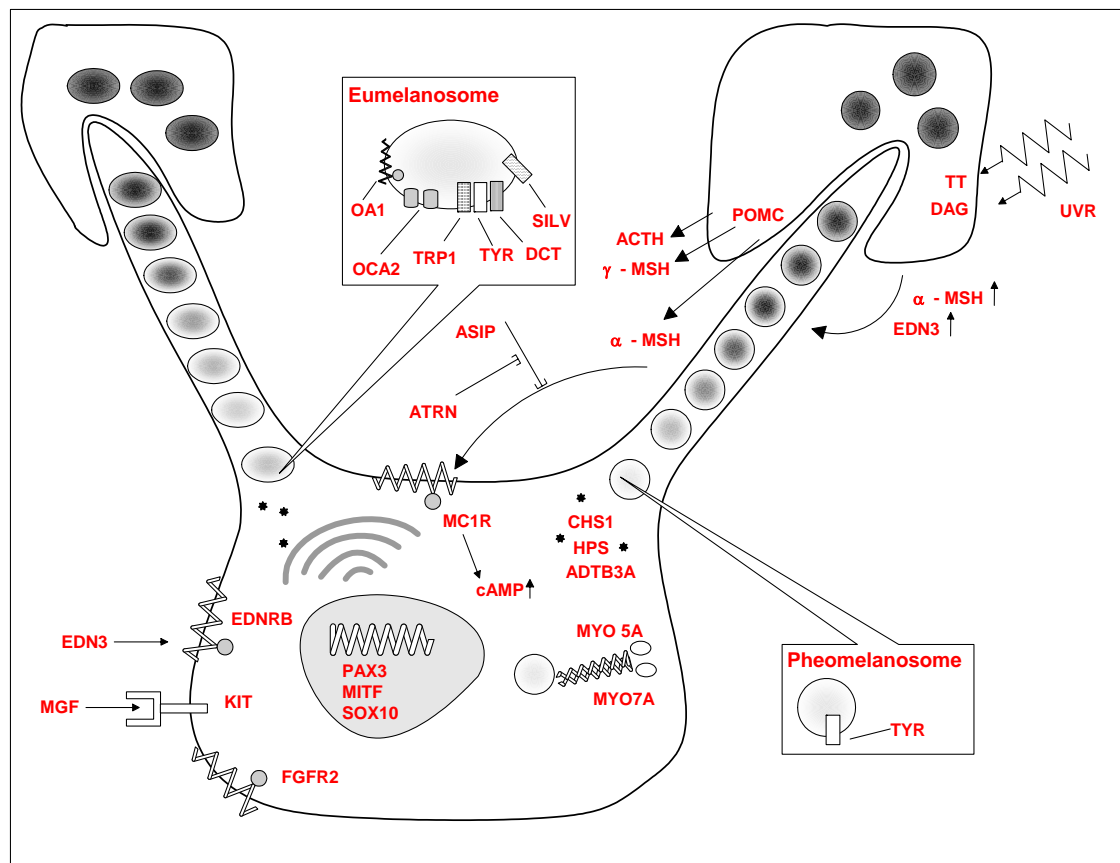


Figure 1.2: Melanocyte-keratinocyte complex. Pigmentation candidate genes are identified in red.

Pigmentation Candidate Genes

By examining the stages involved in melanin biosynthesis, as well as the transport of melanosomes through the melanocyte and into the keratinocyte it is possible to

identify several genes that may affect normal pigmentation variation. However, even before these pathways and cellular structures were well-understood predictions about the number of genes that might be controlling pigmentation phenotype were made. In his 1949 *Principles of Human Genetics* Curt Stern examined data collected on admixed individuals to develop a model that suggested that some combination of four to six pairs of genes with equal and additive effects could explain pigmentation variation between West Africans and northern Europeans (Stern, 1949). Similar conclusions were drawn by Cavalli-Sforza and Bodmer (1971). Computer simulations based on a four-locus model with moderate selection (6%) demonstrate that it is possible for pigmentation phenotype to change from a light extreme to a dark extreme in 10,000 years in keeping with the current evidence from the fossil record (Livingstone, 1964).

Since these early estimates, over fifty pigmentation candidate genes have been identified (<http://albinismdb.med.umn.edu/genes.htm>) primarily through studies of pigmentary disorders or mouse models (Bennett and Lamoreux, 2003). Only six of these genes were examined as part of this work. These genes were selected because of their potential effects on melanin synthesis, melanosome development and morphology, and the switch between the production of either eumelanin or pheomelanin. Other genes, including those believed to be associated with systems such as melanosome transport, are also likely candidates. What follows below is a brief description of the genes considered in this study, as well as the SNPs selected for typing within those genes. It should be noted that associations between a particular SNP and pigmentation phenotype may be indicative that the SNP is actually having a functional effect on pigmentation phenotype, or that it is linked on a haplotype with the functional variant. Such associations should be

taken to indicate that the gene in question affects pigmentation, not necessarily the particular SNP used in the association study.

The gene *TYR* is responsible for the production of the enzyme tyrosinase, the rate limiting enzyme in melanogenesis (Spritz, 1994). Mutations in *TYR* are associated with tyrosinase negative oculocutaneous albinism, or OCA1 (Tomita et al., 1989; Giebel et al., 1991; Oetting et al., 1991; King et al. 1991; Oetting et al., 1993). Three non-albinism related nonsynonymous *TYR* SNPs were typed in this dissertation: A192C, C308G, and A402G. Of these, only A192C has been found to be associated with normal variation in human skin pigmentation, even after controlling for differences in individual admixture proportions in a combined sample of African-Caribbeans and African-Americans (Shriver et al., 2003). However, given the important role that *TYR* plays in melanin production we felt that the two additional nonsynonymous SNPs might also show associations with pigmentation variation.

TYRPI, or tyrosinase-related protein 1, is believed to play a role in stabilizing tyrosinase during melanin production (Jimenez-Cervantes et al. 1998; Kobayashi et al., 1998). Mutations in *TYRPI* have also been associated with OCA3 (Manga et al., 1997). The expression of the product of *TYRPI* has been shown to vary among populations differing in skin pigmentation such that populations with darker skin show higher levels of *TYRPI* expression relative to lighter pigmented populations (Alauf et al., 2003). The SNP typed in *TYRPI*, Gly209Gly, is a synonymous SNP, but was selected because of the possibility that it may be located near a functional SNP in the gene.

Mutations in *OCA2*, the P gene, are associated with tyrosinase positive oculocutaneous albinism, or OCA type 2 (e.g. Durham-Pierre et al., 1994; Lee et al.,

1994; Stevens et al., 1995; Spritz et al., 1995; Suzuki et al., 2003; Yi et al., 2003). Akey et al. (2001) identified interactions between polymorphisms in *OCA2* and the melanocortin 1 receptor (*MC1R*) on skin pigmentation phenotype in a Tibetan population, and Shriver et al. (2003) demonstrated that much like *TYR*, *OCA2* has an effect on normal variation in pigmentation between West Africans and Europeans, even after controlling for the effects of ancestry. The *OCA2* SNP used in the Shriver et al. paper, A355G, was selected to be typed in this study, although it does not result in an amino acid change.

MC1R is a pigmentation candidate gene that shows a high degree of polymorphism outside of Africa (Rana et al., 1999; Harding et al., 2000), although nucleotide diversity in the promoter region is actually highest within Africa (Makova et al., 2001). A number of *MC1R* polymorphisms are associated with red hair, fair skin, and freckling in European populations (e.g. Valverde et al., 1995; Box et al., 1997; Schiöth et al., 1998; Smith et al., 1998; Flanagan et al., 2000; Bastiaens et al., 2001), although variants common outside of Europe have also been described (Yao, et al., 2000; Peng et al., 2001; John et al., 2003). Two *MC1R* polymorphisms were typed in this work, neither of which shows strong association with red hair color (RHC), although Duffy et al. (2004) classified one of these, G92A, as a “weak” RHC allele. However G92A has also been observed at variable frequencies (10-32%) in East and South Asian samples (Rana et al., 1999; Yao et al., 2000; Peng et al., 2001). Given the fact that Austronesian speakers originating from a Southeast Asian homeland migrated through Island Melanesia beginning 3,200 years ago, we felt that G92A might show appreciable variation in the Island Melanesian sample genotyped in Chapter Three. The second *MC1R* SNP typed was a synonymous substitution, G314A. G314A shows a high allele

frequency differences between Africans, Native Americans, and Europeans, and has been used as an ancestry informative allele (AIM) by Shriver et al. (2003) and others. Like G92A, it has also been reported to be polymorphic in East Asians (Yao et al., 2000; Peng et al., 2001).

The agouti signaling protein (*ASIP*) is the antagonist to *MC1R*. Previous studies (Kanetsky et al., 2002) suggested that a polymorphism in the *ASIP* promoter may be associated with dark hair and eyes in European populations. Voisey et al. (2001) did not identify polymorphisms in the *ASIP* coding region among samples of European, African, Native American, or Australian aboriginal descent. I typed a variant in the *ASIP* promoter (A8818G) that was shown during the course of my research to be significantly associated with normal variation in skin pigmentation in African Americans after controlling for the effects of ancestry (Bonilla et al., 2005).

The final pigmentation candidate gene included in this study, membrane associated transport protein (*MATP*), encodes a melanocyte differentiation antigen. *MATP* mutations have been implicated in OCA4 (Newton et al., 2001; Rundshagen et al., 2004; Inagaki et al., 2004). In 2002, polymorphisms in what was then called the *AIM1* (absent in melanoma 1) gene were reported by Nakayama et al. One of these, Phe374Leu, showed large allele frequency differences between populations known to differ in skin pigmentation phenotype. As this work was being completed, Graf et al. (2005) reported significant associations between two *MATP* SNPs (A272G and C374G) and skin, hair, and eye pigmentation in Europeans. Both of the SNPs typed by Graf et al. were included in the genotyping panels used in Chapters Three and Four.

Conclusions

Skin pigmentation is a complex phenotypic trait that has only recently begun to be understood at the genetic level. Although it has been used as a tool to differentiate human populations, pigmentation shows continuous variation across the human species. Many hypotheses suggest that pigmentation differences between populations are due to the effects of natural or sexual selection. The geographic distribution of pigmentation variation is consistent with natural selection-based hypotheses, although sexual selection or genetic drift may also be important (especially in explaining localized variation). This dissertation seeks to explore pigmentation variation from phenotypic, genotypic, and evolutionary perspectives: pigmentation phenotype and genotype are explored in an Island Melanesian sample, while questions regarding the evolution of pigmentation variation are explored using geographically diverse samples using pairwise locus-specific F_{ST} .

Literature Cited

- Agar N, and Young A (2005) Melanogenesis: a photoprotective response to DNA damage? *Mutation Research* 571:121-132.
- Akey J, Wang H, Xiong M, Wu H, Weida L, Shriver M, and Jin L (2001) Interaction between the melanocortin-1 receptor and *P* genes contributes to inter-individual variation in skin pigmentation phenotypes in a Tibetan population. *Human Genetics* 108:516-520.
- Alauf S, Barret K, Bluont M, and Carter N (2003) Ethnic variation in tyrosinase and TYRP1 expression in photoexposed and photoprotected human skin. *Pigment Cell Research* 16:35-42.
- Arsuaga J, Martinez I, Gracia A, Carretero J, and Carbonell E (1993) Three new human skulls from the Sima de los Huesos Middle Pleistocene site in Sierra de Atapuerca, Spain. *Nature* 362:534-537.
- Baker P (1958) Racial differences in heat tolerance. *American Journal of Physical Anthropology* 16:287-305.
- Barbujani G, Magangni A, Minch E, and Cavalli-Sforza L (1997) An apportionment of human genetic diversity. *Proceedings of the National Academy of Sciences* 94:4516-4519.
- Barnicot N (1958) Reflectometry of the skin in southern Nigerians and in some mulattoes. *Human Biology* 30:150-160.

- Bastiaens M, ter Hurne J, Gruis N, Bergman W, Westendorp R, Vermeer B, and Bouwes Bavinck J (2001) The melanocortin-1 receptor gene is the major freckle gene. *Human Molecular Genetics* 10:1701-1708.
- Bennett D, and Lamoreux M (2003) The color loci of mice – a genetic century. *Pigment Cell Research* 16:334-344.
- Blum H (1961) Does the melanin pigment of human skin have adaptive value? *Quarterly Review of Biology* 36:50-63.
- Blumenbach J (1865) *On the Natural Varieties of Mankind*. London: Longman, Green, Longman, Roberts and Green.
- Bonilla C, Boxill L, Donald S, Williams T, Sylvester N, Parra E, Dios S, Norton H, Shriver M, and Kittles R (2005) The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Human Genetics* 116:402-406.
- Boule M (1908) L'Homme fossile de La Chapelle-aux-Saints (Covèze). *Comptes Rendus des Séances de l'Academie des Sciences, Paris* 147:1349-1352.
- Bower C, and Stanley F (1989) Dietary folate as a risk factor for neural tube defects: evidence from a case-control study in Western Australia. *Medical Journal of Australia* 150:613-619.
- Box N, Wyeth J, O'Gorman L, Martin N, and Sturm R (1997) Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. *Human Molecular Genetics* 6:1891-1897.

- Branda R, and Eaton J (1978) Skin color and nutrient photolysis: an evolutionary hypothesis. *Science* 201:625-626.
- Byard P, and Lees F (1982) Skin colorimetry in Belize. II. Inter- and Intra-population variation. *American Journal of Physical Anthropology* 58:215-219.
- Cann R, Stoneking M, and Wilson A (1987) Mitochondrial DNA and human evolution. *Nature* 325:31-36.
- Cavalli-Sforza L, Menozzi P, and Piazza A (1994) *The History and Geography of Human Genes*. Princeton: Princeton University Press.
- Cavalli-Sforza L, and Bodmer W (1971) *The Genetics of Human Populations*. San Francisco: W.H. Freeman and Company, pp. 527-531.
- Clark J, Beyene Y, WoldeGabriel G, Hart W, Renne P, Gilbert H, Defleur A, Suwa G, Katoh S, Ludwig K, Boisserie J, Asfaw B, and White T (2003) Stratigraphic, chronological and behavioural contexts of Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423:747-752.
- Conway D, and Baker P (1972) Skin reflectance of Quechua Indians: the effects of genetic admixture, sex, and age. *American Journal of Physical Anthropology* 36:367-282.
- Cowles R (1959) Some ecological factors bearing on the origin and evolution of pigment in the human skin. *American Naturalist* 93:283-293.
- Darwin C (1871) *The Descent of Man, and Selection in Relation to Sex*. London: J. Murray.
- deMenocal P (1995) Plio-Pleistocene African Climate. *Science* 270:53-59.

- Diamond J (1992) *The Third Chimpanzee*. New York: Harper Collins Publishers.
- Dmiel R, Prevoulotzky A, and Shkolnik A (1980) Is a black coat in the desert a means of saving metabolic energy? *Nature* 283:761-762.
- Dubois E (1926) On the principle characters of the cranium and the brain, the mandible and the teeth of *Pithecanthropus erectus*. *Proceedings of the Academy of Science Amsterdam* 27:265-278.
- Duffy D, Box N, Chen W, Palmer J, Montgomery G, James M, Hayward N, Martin N, and Sturm R (2004) Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. *Human Molecular Genetics* 13:447-461.
- Durham-Pierre D, Gardner J, Nakatsu Y, King R, Francke U, Ching A, Aquaron R, del Marmol V, and Brilliant M (1994) African origin of an intragenic deletion of the human P gene in tyrosinase positive oculocutaneous albinism. *Nature Genetics* 7:176-179.
- Flanagan N, Healy E, Ray A, Philips S, Todd C, Jackson I, Birch-Machin M, and Rees J (2000) Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation. *Human Molecular Genetics* 9:2531-2537.
- Flemming A, and Copp A (1998) Embryonic folate metabolism and mouse neural tube defects. *Science* 280:2107-2109.
- Friedmann P, and Gilchrest B (1987) Ultraviolet radiation directly induced pigment production by cultured human melanocytes. *Journal of Cellular Physiology* 133:88-94.

- Frost P (1988) Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine* 32:38-58.
- Frost P (1994) Geographic distribution of human skin colour: a selective compromise between natural selection and sexual selection? *Human Evolution* 9:141-153.
- Gabunia L, and Vekua A (1995) A Plio-Pleistocene hominid from Dmanisi, East Georgia, Caucasus. *Science* 288:1019.
- Giebel L, Musarella M, and Spritz R (1991) A nonsense mutation in the tyrosinase gene of Afghan patients with tyrosinase negative (type IA) oculocutaneous albinism. *Journal of Medical Genetics* 28:461-467.
- Goldstein D, Ruiz Linares A, Cavalli-Sforza L, and Feldman M (1995) Genetic absolute dating based on microsatellites and the origin of modern humans. *Proceedings of the National Academy of Sciences, USA* 92:6723-6727.
- Gould S (1981) *The Mismeasure of Man*: W.W. Norton and Company.
- Graf J, Hodgson R, and van Daal A (2005) Single nucleotide polymorphisms in the *MATP* gene are associated with normal human pigmentation variation. *Human Mutation* 25:278-284.
- Haile-Selassie Y, Asfaw B, and White T (2004) Hominid cranial remains from upper Pleistocene deposits at Aduma, Middle Awash, Ethiopia. *American Journal of Physical Anthropology* 123:1-10.
- Harding R, Healy E, Ray A, Ellis N, Flanagan N, Todd C, Dixon C, Sajantila A, Jackson I, Birch-Machin M, and Rees J (2000) Evidence for variable selective pressures at MC1R. *American Journal of Human Genetics* 66:1351-1361.

- Harsanyi Z, Post P, Brinkmann J, Chedekel M, and Diebel R (1980) Mutagenicity of melanin from human red hair. *Experientia* 36:291-292.
- Harvey R (1985) Ecological factors in skin color variation among Papua New Guineans. *American Journal of Physical Anthropology* 66:407-416.
- Herodotus (415 BC) *The Histories*. New York: Penguin Books.
- Holick M, Smith E, and Pincus S (1987) Skin as the site of vitamin D synthesis and target tissue for 1,25-dihydroxyvitamin D₃. *Archives of Dermatology* 123:1677-1683.
- Holubar K (1996) What is a caucasian? *Journal of Investigative Dermatology* 106:800.
<http://albinismdb.med.umn.edu/genes.htm>.
- Inagaki K, Suzuki T, Shimizu H, Ishii N, Umezawa Y, Tada J, Kikuchi N, Takata M, Takamori K, Kishibe M, Tanaka M, Miyamura Y, Ito S, and Tomita Y (2004) Oculocutaneous albinism type 4 is one of the most common types of albinism in Japan. *American Journal of Human Genetics* 74:466-471.
- Jablonski N, and Chaplin G (2000) The evolution of human skin coloration. *Human Evolution* 39:57-106.
- Jimenez-Cervantes C, Martinez-Esparza F, Solano J, Lozan J, and Garcia-Borron (1998) Molecular interactions within the melanogenic complex: formation of heterodimers of tyrosinase and TRP1 from B16 mouse melanoma. *Biochem Biophys Res Commun* 253:761-767.

- John P, Makova K, Li W, Jenkins T, and Ramsay M (2003) DNA polymorphism and selection at the melanocortin-1 receptor gene in normally pigmented southern African individuals. *Annals of the NY Academy of Sciences* 994.
- Jorde L, Watkins W, Bamshad M, Dixon M, Ricker C, Seielstad M, and Batzer M (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *The American Journal of Human Genetics* 66:979-988.
- Kahane P (1967) *Ancient and Classical Art*. New York: Dell.
- Keeler C (1953) The Caribe Moon Child and its Heredity. *Journal of Heredity* 44:162-171.
- Keeler C (1968) Note on sweating and odor of Cuna albinos. *Dermatologia Internationalis* 7:78-80.
- King R, Mentink M, and Oetting W (1991) Non-random distribution of missense mutations within the human tyrosinase gene in type I (tyrosinase-related) oculocutaneous albinism. *Molecular Biology and Medicine* 8:19-29.
- Kittler R, Kayser M, and Stoneking M (2003) Molecular evolution of *Pediculus humanus* and the origin of clothing. *Current Biology* 13:1414-1417.
- Kobayashi T, Imokawa G, Bennett D, and Hearing V (1998) Tyrosinase stabilization by Tyrp1 (the brown locus). *Journal of Biological Chemistry* 273:31801-31805.

- Koch W, and Chedekel M (1986) Photoinitiated DNA damage by melanogenic intermediates in vitro. *Photochemistry and Photobiology* 44:703-710.
- Ladell W (1964) Terrestrial animals in humid heat. In D Dill (ed.): *Handbook of Physiology: Adaptation to the Environment*. Baltimore: Williams and Wilkins, pp. 625-659.
- Laurence K, James N, Miller M, Tennant G, and Campbell H (1981) Double-blind randomised controlled trial of folate treatment before conception to prevent recurrence of neural-tube defects. *British Medical Journal (Clinical Research Edition)* 282:1509-1511.
- Lee S, Nicholls R, Schnur R, Guida L, Lu-Kuo J, Spinner N, Zackai E, and Spritz R (1994) Diverse mutations of the P gene among African-Americans with type II (tyrosinase positive) oculocutaneous albinism (OCA2). *Human Molecular Genetics* 3:2047-2051.
- Lewontin R (1972) The apportionment of human diversity. In T Dobzhansky, M Hecht and W Steere (eds.): *Evolutionary Biology* 6. New York: Appleton Century Crofts, pp. 381-398.
- Linneaus C (1735) *Systema Naturae*. Nieuwkoop: B. de Graf.
- Livingstone F (1964) Polygenic models for the evolution of human skin color differences. *Human Biology* 41:480-493.
- Lock-Andersen J, and Wulf H (1997) Seasonal variation of skin pigmentation. *Acta Derm Venereol* 77:219-221.

- Loomis W (1967) Skin pigment regulation of vitamin-D biosynthesis in man. *Science* 157:501-506.
- Luande J, Henschke C, and Mohamed N (1985) The Tanzanian human albino skin: natural history. *Cancer* 55:1823-1828.
- Mackintosh J (2001) The antimicrobial properties of melanocytes, melanosomes, and melanin and the evolution of black skin. *Journal of Theoretical Biology* 211:101-113.
- Makova K, Ramsay M, Jenkins T, and Li W (2001) Human DNA sequence variation in a 6.6-kb region containing the melanocortin 1 receptor promoter. *Genetics* 158:1253-1268.
- Manga P, Kromberg J, Box N, Sturm R, Jenkins T, and M R (1997) Rufous oculocutaneous albinism in southern African Blacks is caused by mutations in the TYRP1 gene. *American Journal of Human Genetics* 61:1095-1101.
- Marcy S (1839) On the albino. *American Journal of Medical Science* 24:517-518.
- Mathur U, Datta S, and Mathur B (1977) The effect of aminopterin-induced folic acid deficiency on spermatogenesis. *Fertility and Sterility* 28:1356-1360.
- McMichael A, and Hall A (1997) Does immunosuppressive ultraviolet radiation explain the latitude gradient for multiple sclerosis? *Epidemiology* 8:642-645.
- Miller M (1986) *The Murals of Bonampak*. Princeton: Princeton University Press.
- Montagna W, and Carlisle K (1991) The architecture of black and white facial skin. *Journal of the American Academy of Dermatologists* 24:929-937.

- Murray F (1934) Pigmentation, sunlight and nutritional disease. *American Anthropologist* 36:438-445.
- MVSR (1991) Prevention of neural tube defects: Results of the Medical Research Council Vitamin Study. *Lancet* 338:131-137.
- Nakayama K, Fukamachi S, Kimura H, Koda Y, Soemantri A, and Ishida T (2002) Distinctive distribution of AIM1 polymorphism among major human populations with different skin color. *Journal of Human Genetics* 47:92-94.
- Newton J, Cohen-Barak O, Hagiwara N, Gardner J, Davisson M, King R, and Brilliant M (2001) Mutations in the human orthologue of the mouse underwhite gene (*uw*) underlie a new form of oculocutaneous albinism, OCA4. *American Journal of Human Genetics* 69:981-988.
- Nordlund J, Boissy R, Hearing V, King R, and Ortonne J (1998) *The Pigmentary System - Physiology and Pathophysiology*. Oxford: Oxford University Press.
- Oetting W, and King R (1993) Molecular basis of type I (tyrosinase-related) oculocutaneous albinism: mutations and polymorphisms of the human tyrosinase gene. *Human Mutation* 2:1-6.
- Oetting W, Mentink M, Summers C, Lewis R, White J, and King R (1991) Three different frameshift mutations of the tyrosinase gene in type I-A (tyrosinase-negative) oculocutaneous albinism in Puerto Rico. *American Journal of Human Genetics* 52:17-23.

- Ohmart R, and Lasiewski R (1971) Roadrunners: energy conservation by hypothermia and absorption of sunlight. *Science* 172:67-69.
- Okoro A (1975) Albinism in Nigeria: a clinical and social study. *British Journal of Dermatology* 92:485-492.
- Pathak M, and Fitzpatrick T (1974) The role of natural photoprotective agents in human skin. In M Pathak (ed.): *Sunlight and Man*. Tokyo: University of Tokyo Press, pp. 725-750.
- Peng S, Lu X, Luo H, Xinag-Yu J, and Zhang Y (2001) Melanocortin-1 receptor gene variants in four Chinese ethnic populations. *Cell Research* 11:81-84.
- Pliny (77 AD) *Natural History*, Book Seven.
- Post P, Daniels F, and Binford R (1975) Cold injury and the evolution of 'white' skin. *Human Biology* 39:131-143.
- Post P, and Rao D (1977) Genetic and environmental determinants of skin color. *American Journal of Physical Anthropology* 47:399-402.
- Ptolemy (140 AD) *Tetrabiblos*.
- Rak Y, Kimbel W, and Hovers E (1994) A Neandertal infant from Amud Cave, Israel. *Journal of Human Evolution* 26:313-324.
- Rana B, Hewett-Emmett D, Jin L, Chang B, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde L, Ramsay M, Jenkins T, and Li W (1999) High polymorphism at the human melanocortin 1 receptor locus. *Genetics* 151:1547-1557.
- Riggs S, and Sargent F (1964) Physiological regulation in moist heat by young American Negro and white males. *Human Biology* 36:339-353.

- Rightmire G, and Deacon H (1991) Comparative studies of Late Pleistocene human remains from Klasies River Mouth, South Africa. *Journal of Human Evolution* 20:131-156.
- Roberts D, and Kahlon D (1976) Environmental correlations of skin colour. *Annals of Human Biology* 3:11-22.
- Robins A (1991) *Biological perspectives on human pigmentation*. New York: Cambridge University Press.
- Rogers A, Iltis D, and Wooding S (2004) Genetic variation at the MC1R locus and the time since loss of human body hair. *Current Anthropology* 45:105-108.
- Rundshagen U, Zuhlke C, Opitz S, Schwinger E, and Kasmann-Kellner B (2004) Mutations in the MATP gene in five German patients affected by oculocutaneous albinism type 4. *Human Mutation* 23:106-110.
- Rush B (1799) Observations intended to favour a supposition that the black color (as it is called) of the Negroes is derived from leprosy. *American Philosophical Society Transactions* 4:289-297.
- Sartono S (1971) Observations on a new skull of *Pithecanthropus erectus* (*Pithecanthropus VIII*) from Sangiran, Central Java. *Proceedings of the Academy of Science, Amsterdam* 74:185-194.
- Schacter B, Lederman M, LeVine M, and Ellner J (1983) Ultraviolet radiation inhibits human natural killer activity and lymphocyte proliferation. *Journal of Immunology* 130:2484-2487.

- Schiöth H, Philips S, Rudzish R, Birch-Machin M, Wikberg J, and Rees J (1999) Loss of function mutations of the human melanocortin 1 receptor are common and are associated with red hair. *Biochem Biophys Res Commun* 260:488-491.
- Schmitz S, Thomas P, Allen T, Poznansky M, and Jimbow K (1995) Dual role of melanins and melanin precursors as photoprotective and phototoxic agents: inhibition of ultraviolet radiation-induced lipid peroxidation. *Photochemistry and Photobiology* 61:650-655.
- Seielstad M, Bekele E, Ibrahim M, Toure A, and M T (1999) A view of modern human origins from Y chromosome microsatellite variation. *Genome Research* 9:558-567.
- Shriver M, Parra E, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Human Genetics* 112:387-399.
- Singer R, and Wymer J (1982) *The Middle Stone Age at Klasies River Mouth in South Africa*. Chicago: University of Chicago Press.
- Smith R, Healy E, Siddiqui S, Falanagan N, Steijlen P, Rosdahl I, and Japques J (1998) Melanocortin 1 receptor variants in Irish populations. *Journal of Investigative Dermatology* 111:119-122.
- Smith S (1788) *An Essay on the causes of the variety of complexion and figure in the human species*. Edinburgh: C. Elliot.

- Snowden F (1970) Blacks in Antiquity: Ethiopians in the Greco-Roman Experience. Cambridge, MA: The Belknap Press of Harvard University Press.
- Spritz R (1994) Molecular genetics of oculocutaneous albinism. *Human Molecular Genetics* 3:1469-1475.
- Spritz R, Fukai K, Holmes S, and Luande J (1995) Frequent intragenic deletion of the P gene in Tanzanian patients with tyep II oculocutaneous albinism (OCA2). *American Journal of Human Genetics* 56:1320-1323.
- Staples J, Ponsonby A, Lim L, and McMichael A (2003) Ecologic analysis of some immune-related disorders, including type 1 diabetes, in Australia: latitude, regional ultraviolet radiation, and disease prevalence. *Environmental Health Perspectives* 111:518-523.
- Stern C (1949) Polygenic Inheritance. In C Stern (ed.): *Principles of Human Genetics*. San Francisco: W.H. Freeman and Company.
- Stevens G, van Beukering J, Jenkins T, and Ramsay M (1995) An intragenic deletion of the P gene is the common mutation causing tyrosinase-positive oculocutaneous albinism in Southern African Negroids. *American Journal of Human Genetics* 56:586-591.
- Stoneking M, Fontius J, Clifford S, Soodyall H, Arcot S, Nilmani S, Jenkins T, Tahir M, Deininger P, and Batzer M (1997) Alu insertion polymorphisms and human evolution: evidence for a larger population size in Africa. *Genome Research* 7:1061-1071.

- Stoneking M, Sherry S, Redd A, and Vigilant L (1992) New approaches to dating suggest a recent age for the human mtDNA ancestor. *Philos Trans R Soc Lond B Biol Sci* 337:167-175.
- Strydom N, and Wyndham C (1963) Natural state of heat acclimatization of different ethnic groups. *Federation Proceedings* 22:801-809.
- Sturn R (1998) Human pigmentation genes and their response to solar UV radiation. *Mutation Research* 422.
- Sturn R, Teasdale R, and Box N (2001) Human pigmentation genes: identification, structure and consequences of polymorphic variation. *Genes and Genetic Systems* 277:49-62.
- Suzuki T, Miyamura Y, Matsunaga J, Shimizu H, Kawachi Y, Ohyama N, Ishikawa O, Ishikawa T, Terao H, and Tomita Y (2003) Six novel P gene mutations and oculocutaneous albinism type 2 frequency in Japanese albino patients. *Journal of Investigative Dermatology* 120.
- Termorshuizen F, Garssen J, Norval M, Koulu L, Laihia J, Leino L, Jansen C, De Gruijl F, Gibbs N, De Simone C, and Van Loveren H (2002) A review of studies on the effects of ultraviolet irradiation on the resistance to infections: evidence from rodent infection models and verification by experimental and observational human studies. *International Immunopharmacology* 2:263-275.
- Thomson R, Pritchard J, Shen P, Oefner P, and Feldman M (2000) Recent common ancestor of human Y chromosomes: Evidence from DNA sequence data. *Proceedings of the National Academy of Sciences, USA* 97.

- Tishkoff S, and Verrelli B (2003) Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annual Review of Genomics and Human Genetics* 4:293-340.
- Tobias P (1961) Studies on skin reflectance in Bushmen-European hybrids. *Proceedings, Second International Congress of Human Genetics, Rome*:461-471.
- Tomita Y, Takeda A, Okinaga S, Tagami H, and Shibahara S (1989) Human oculocutaneous albinism caused by single base insertion in the tyrosinase gene. *Biochemical and Biophysical Research Communications* 164.
- Underhill P, Passarino G, Lin A, Shen P, Mirazon Lahr M, Foley R, Oefner P, and Cavalli-Sforza L (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of Human Genetics* 65:43-62.
- Valverde P, Healy E, Jackson I, Rees J, and Thody A (1995) Variants of the melanocyte stimulating hormone receptor gene are associated with red hair and fair skin. *Nature Genetics* 11:328-330.
- Van den Berghe P, and Frost P (1986) Skin color preference, sexual dimorphism and sexual selection: a case of gene culture co-evolution? *Ethnic and Racial Studies* 9:87-113.
- Vekua A, Lordkipanidze D, Rightmire G, Agusti J, Ferring R, Maisuradze G, Mouskhelishvili A, Nioradze M, Ponce de Leon M, Tappen M, Tvalchredlidze M, and Zollikofer C (2002) A new skull of early Homo from *Dmanisi*, Georgia. *Science* 297:85-89.

- Vigilant L, Stoneking M, Harpending H, Hawkes K, and Wilson A (1991) African populations and the evolution of human mitochondrial DNA. *Science* 253:1503-1507.
- Voisey J, Box N, and van Daal A (2001) A polymorphism study of the human Agouti gene and its association with MC1R. *Pigment Cell Research* 14:264-267.
- Walter H (1971) Remarks on the environmental adaptation of man. *Humangenetik* 13.
- Wasserman H (1965) Human pigmentation and environmental adaptation. *Archives of Environmental Health* 11:691-694.
- Watkins W, Rogers A, Ostler C, Wooding S, Bamshad M, Brassington A, Carroll M, Nguyen S, Walker J, Prasad B, Reddy P, Das P, Batzer M, and Jorde L (2003) Genetic variation among world populations: inferences from 100 Alu insertion polymorphisms. *Genome Research* 13:1607-1618.
- Weitzen M, and Bonavida B (1984) Mechanism of inhibition of human natural killer activity by ultraviolet radiation. *Journal of Immunology* 133:3128-3132.
- Westbroek W, Lambert J, and Naeyaert J (2001) The dilute locus and Griscelli syndrome: gateways towards a better understanding of melanosome transport. *Pigment Cell Research* 14:320-327.
- White T, Asfaw B, DeGusta D, Gilbert H, Richards G, Suwa G, and Howell F (2003) Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423:742-747.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15:323-354.

Yao Y, Lu X, Luo H, Li W, and Zhang Y (2000) Gene admixture in the silk road region of China: evidence from mtDNA and melanocortin 1 receptor polymorphism. *Genes and Genetic Systems* 75:173-178.

Yi Z, Garrison N, Cohen-Barak O, Karafet T, King R, Erickson R, Hammer M, and Brilliant M (2003) A 122.5-kilobase deletion of the P gene underlies the high prevalence of oculocutaneous albinism type 2 in the Navajo population. *American Journal of Human Genetics* 72:62-72.

Chapter 2

Skin and Hair Pigmentation Variation in Island Melanesia

Introduction

The human species shows remarkable variation in both skin and hair pigmentation. Both traits, particularly skin pigmentation, have commonly been used as tools to classify individuals into discrete racial groups. However, a closer inspection has shown that humans cannot be so neatly categorized. Skin pigmentation, despite often being synonymous with racial taxonomies, shows substantial variation within human populations and within regional groupings or “races.” Pigmentation variation within and among populations has been explained as the result of either natural selection (e.g. Cowles, 1959; Wasserman, 1965; Loomis, 1967; Walter, 1971; Post et al., 1975; Branda and Eaton, 1978; Jablonski and Chaplin, 2000; Mackintosh, 2001) or sexual selection (Darwin, 1871; Diamond, 1988; Frost, 1988; Diamond, 1992). The selection-based hypotheses have focused on the relationship between skin pigmentation and ultraviolet radiation, or UVR (Walter, 1971; and as explored in Relethford, 1997; Jablonski and Chaplin, 2000; Chaplin, 2004). High levels of UVR can lead to skin damage such as sunburn or skin cancer, as well as the breakdown of folic acid (Branda and Eaton, 1978). As melanin provides some protection from UVR-induced damage (e.g. Pathak and Fitzpatrick, 1974; Kollias et al., 1991; Sheehan et al., 2002), a compelling selectionist argument is that darkly melanized skin is an adaptation in regions of high UVR. While photodamage and nutrient photolysis may not be very strong selective pressures in regions where UVR is low, these regions present their own challenges. Specifically, it has been suggested that lighter pigmentation in these regions may have been an adaptation to ensure the production of sufficient levels of Vitamin D₃ in the skin (Murray, 1934; Loomis, 1967). This has been a commonly accepted explanation for the general

trend towards lighter skin pigmentation in low UVR regions, and recent work by Jablonski and Chaplin (2000) strengthens the case for it.

This paper examines pigmentation variation in Island Melanesia, an equatorial area renowned for its biological and linguistic diversity (See Figure 2.1). The peopling of Island Melanesia and the larger region of Oceania has been studied from linguistic, archaeological, and biological perspectives. While each of these approaches has revealed particular aspects of the history of human migrations in this region, a synthesis has begun to emerge concerning the outlines of modern human dispersals in Island Melanesia. Human populations had expanded across the Old World and into Australia by 50,000 BP (Roberts et al., 1994; Bowler et al., 2003). The earliest human presence in New Guinea can be dated to at least 40,000 BP, while recent estimates for the earliest occupation of the Bismarck Archipelago (namely New Ireland) also occur around 40,000 years BP (Leavesley et al., 2002), implying sailing capabilities at that time. These dates for human occupation at New Ireland suggest an even earlier settlement date for New Guinea, possibly contemporaneous with Australian settlement (Allen, 2003). Evidence for human habitation on the interior of New Britain dates to 35,000 BP (Pavrides and Gosden, 1994), and by 29,000 BP humans had reached , Buka and the North Solomons (Wickler and Spriggs, 1988). By 20,000 years ago, archaeological evidence indicates the intentional introduction of animals into the region as well as more extensive long range interactions or trade (e.g. Summerhayes and Allen, 1993; Allen 1996; Leavesley and Allen, 1998). A very distinctive cultural horizon in the Bismarcks dates to about 3,200 years ago, called Lapita, after its distinctive pottery style. It was marked by a number of other novel or intrusive material culture traits, including a number of domesticates (pigs, chickens, dogs), unique shell ornament types, extensive trading networks, and new settlement patterns, focusing on certain shore locations and

peripheral islands (Spriggs, 1995). It has been linked to the movement of a new language into the region (Proto-Oceanic, an Austronesian language). This was the ancestor of all the Austronesian languages spoken today throughout Island Melanesia, Polynesia and most of Micronesia. However, many languages that belong to a non-Austronesian stratum survive, particularly in the interiors of New Guinea and the larger islands of Island Melanesia. Spriggs (1997) has hypothesized that the Lapita culture, its extensive trade networks, and the chain of Austronesian dialects that developed with it helped to homogenize Island Melanesia during the period of 2500-2000 BC., and that subsequently there was a contraction of trade networks and a local diversification of Austronesian-speaking populations in the region. As a corollary to this model, the isolation accompanying these changes would be expected to increase, or at least maintain, genetic diversity across the region.

Earlier genetic studies tended to focus on whether Austronesian (Oceanic) speaking populations could be distinguished from non-Austronesian (Papuan) ones, or whether simple geographic propinquity in this small region was the best discriminator. The results have been conflicting, although the degree of variation is very high (Friedlaender, 1975; Friedlaender, 1987; Merriwether, Friedlaender et al., 1999; Robledo et al., 2004; Friedlaender et al., in press; *inter alia*).

As reported here, pigmentation variation in this region echoes the remarkable genetic diversity observed in previous studies. Although the inhabitants of Island Melanesia are generally darkly pigmented, their skin pigmentation shows a surprising amount of diversity across the region. The explanation for the pattern of diversity within this region cannot be tied directly to contemporary UVR exposure, but rest instead on the history of ancient population migrations or other factors.

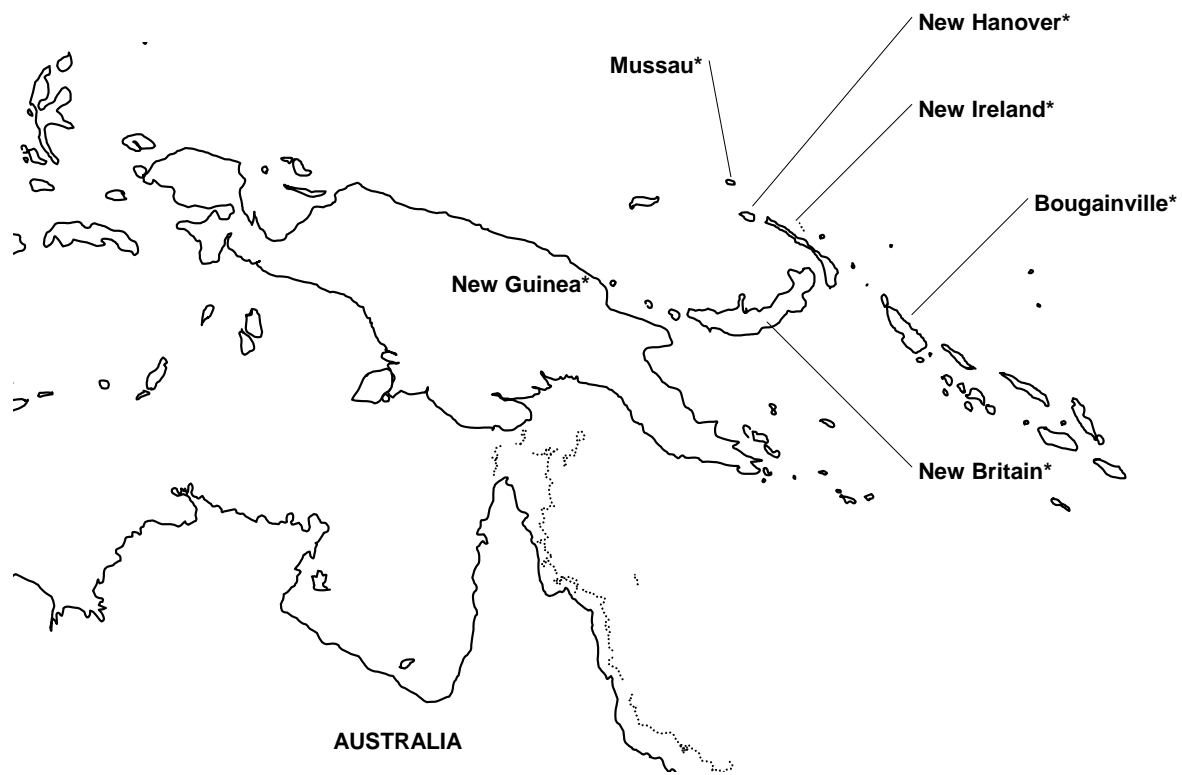


Figure 2.1: Map of the study region in Island Melanesia. Samples were obtained from the islands of New Guinea, New Britain, New Ireland, Mussau, New Hanover, and Bougainville (denoted *).

Materials and Methods

One thousand one hundred and thirty-five adult individuals from 12 Austronesian and 6 Papuan speaking groups from 6 islands in Island Melanesia were measured for skin and hair pigmentation by HLN and JSF. These individuals were measured on New Britain, New Ireland, and New Hanover in the Bismarck Archipelago, and near-by Bougainville. Both Austronesian (AN) and Papuan (P) language speaking individuals were included. This work was carried out as

part of a larger project on the population history and settlement of Island Melanesia. As such, information concerning each individual's age, village of origin, language, and the villages and languages of their parents was also collected along with a 10 cc blood sample for genetic analysis. All individuals involved in this study gave their informed consent to participate as research subjects. IRB approval for the pigmentation aspects of this work was obtained from Penn State University (IRB # 00M558-2), Temple University (IRB # 99-226) and The Papua New Guinea Medical Research Advisory Committee.

Measurement

Measurements of skin and hair pigmentation were taken using the DermaSpectrometer (Cortex Technology, Hadsund, Denmark), a narrow band reflectance spectrophotometer. The DermaSpectrometer is designed to estimate the concentrations of hemoglobin and melanin, the primary chromophores of the skin, based on the work of Diffey et al. (1984). Both hemoglobin and melanin absorb light at lower wavelengths, with hemoglobin showing peak absorbance in the shorter (green) wavelengths, followed by a sharp drop off in absorbance in the longer (red) wavelengths (above 650nm). Melanin, however, absorbs consistently across the visible spectrum. Based on these differences in absorbance, Diffey et al. (1984) proposed that the reflectance of narrow-band light in the red spectrum would result in an estimate of the melanin content of an individual's skin (M index), using the following equation:

$$M = \log_{10} (1/\% \text{ red reflectance})$$

The M index is useful in studies of pigmentation variation because it measures the amount of skin pigmentation that is due primarily to the effects of melanin with limited confounding effects from hemoglobin.

Many anthropological studies of skin pigmentation variation have used reflectometers such as the E.E.L. (Evans Electro Selenium Limited) or Photovolt line of instruments (UMM instruments). These reflectometers measure percent reflectance at different wavelengths by passing light through differently colored filters. For example, skin pigmentation studies were commonly performed using red, green, and blue filters (425, 545, and 685 nm for the E.E.L. instrument). Percent reflectance in the red filter (685 nm) is the closest approximation to the Dermapsectrometer's M index, but it should be noted that the results are not directly comparable.

Prior to use the DermaSpectrometer was calibrated using the white and black standard provided by the manufacturer and in accordance with recommended practices. Three measurements were taken on the upper inner left and right arms of subjects, following the recommendations of Shriver and Parra (2000). The upper inner arm was selected as the measurement site because it is a region of the body that is generally unexposed to UVR, allowing for a more accurate measurement of constitutive (unexposed) rather than facultative (tanned) skin pigmentation. Three measurements were also taken of the hair at the crown. Cases where hair had been bleached or colored, or was gray or thin, were noted and subsequently excluded from later analysis. The three hair measurements were averaged together to give a mean hair M index value per individual. The six skin measurements were averaged to yield mean skin M index values for each individual. For analysis of pigmentation variation subjects were classified into categories according to sex, linguistic phylum (Austronesian or Papuan), island, and neighborhood.

Island Classification

Individuals were placed into categories according to their geographic region of origin in Island Melanesia. An individual was placed into a particular island category if both he or she and his or her parents were from that island. In some cases a subject's reported language was used to help determine island placement. For example, recently a group of Kapugu-speakers from the island of Mussau migrated to Kavieng town in New Ireland. Individuals measured in Kavieng who identified themselves and their parents as being Kapugu speakers but listing Kavieng as their home were assigned to Mussau Island rather than to New Ireland. Individuals with parents from different islands were excluded from island-level analyses. In this way, a total of 1046 individuals were assigned to the following islands: Bougainville (153), New Hanover (102), Mussau (35), Manus (2), New Britain (491), New Ireland (242), and New Guinea (21). Due to the extremely small number of individuals that could be assigned to the island of Manus it was excluded from further island-level analyses.

Linguistic Classification

As mentioned, considerable work in the Southwest Pacific has explored the relation between genetic and linguistic variation. The work of Giles et al. (1965), which identified clear genetic distinctions between Austronesian and Papuan language speakers in the Markham Valley on the island of New Guinea, led to investigations of Island Melanesian diversity according to linguistic affiliation (primarily Austronesian vs. Papuan language speakers). While later work (e.g. Serjeantson and Gao, 1995; Merriwether, Friedlaender et al., 1999) did not find genetic variation following the clear

linguistic divisions observed in the Markham Valley, variation in regions such as Bougainville could be seen to be at least partly (but certainly not entirely) related to linguistic affiliation. Specifically, in Bougainville linguistic distinctions could explain anthropometric variation better than geographic distance alone. The relationship between language and this biological variation may be explained by past isolation between two groups, due to either different arrival times to the region or isolation due to small population size (Friedlaender, 1975; Friedlaender, 1987).

To determine if pigmentation also showed an association with linguistic affiliation, individuals were classed as Austronesian (AN) or Papuan (P), acknowledging that the Papuan languages are extremely diverse and do not form a clearly related family. An individual and both parents had to speak languages belonging to the same linguistic phylum to be classified as either P or AN. As with the island classification scheme, individuals with parents speaking languages belonging to different phyla were excluded from AN/P analyses. Table 2.1 shows the listing of AN and P language groups included in this study, and the number of individuals in each group. The undivided categories of “Austronesian” and Papuan” are comprised of individuals belonging to AN or P language groups where the population $n < 10$, as well as those individuals whose parents spoke different languages from the same phylum. We have classified Madak of New Ireland as Papuan rather than Austronesian, because Ross (1994) argued it was originally a Papuan language with a subsequent Austronesian overlay. We compared pigmentation between Austronesian and Papuan speakers across the region, as well as specifically within the separate large islands.

Table 2.1: Sample composition of Austronesian and Papuan speakers

Austronesian	n	Papuan	n
Kapugu	34	Aita	36
Kove	45	Anem	55
Kuanua	18	Ata	68
Lavongai	102	Kuot	53
Mamusi	80	Madak	26
Melamela	37	Sepik	11
Nailik	27	"Papuan"	29
Nakanai	115		
Notsi	20		
Saposa	42		
Teop	23		
Tigak	27		
"Austronesian"	132		
Total	702		278

Neighborhood Classification

It is important to make our sampling rationale explicit. Our objective was to capture traditional patterns of genetic variation in rural regions of the major islands. Therefore, we selected a set of villages that were distributed across major linguistic groups and also across island sections. Of course, accessibility by road or boat and group interest in participation were important factors. In practice, this generally meant focusing on the few Papuan speaking populations in the region and their immediate Austronesian-speaking neighbors. We also tried to sample at least two villages from different areas within each language group, in order to obtain some measure of within-language group variation. We refer to our resulting samples as coming from different neighborhoods – neither strictly linguistic nor geographically determined, but a combination of the two.

This is a modification of the more intensive sampling strategy used by Friedlaender in his much earlier survey in a section of Bougainville (1975), where 18 villages were sampled along a

single path that crossed a number of different language groups. In that study, marital migration distances were found to be extremely low for both men and women (typically less than 1 km). Language boundaries appeared to have a very slight additional effect on restricting contemporary marital migration rates, but some biological differences among the villages still reflected linguistic distinctions. Friedlaender argued this was because language affinity was often a strong indication of long standing population relationship. We decided this justified using a combination of geographic and linguistic considerations in our sampling of neighborhoods.

Neighborhood classification was carried out in a manner similar to that for linguistic phyla and region: an individual and both of his or her parents needed to be from the same neighborhood (i.e. same language and residential area) in order to be assigned to that group. In this manner 829 individuals could be classified into 65 neighborhoods. However some of these neighborhoods were not used because of insufficient numbers, and we included only those populations for which 10 or more individuals were present in our sample (resulting in 28 neighborhoods, totaling 796 individuals). In 306 cases, an individual and his or her parents identified with different neighborhood groups (implying some difference in language and residential area). These individuals were excluded from neighborhood-level analysis.

Latitude/longitude Classification

Many hypotheses about the global distribution of skin pigmentation variation have focused on the relationship between pigmentation and UVR. Until recently it was difficult to directly measure UVR, and latitude was used as a proxy (Roberts and Kahlon, 1976; Tasa et al., 1985; Relethford, 1997). In an effort to explore this relationship, as well as to explore general geographical trends in pigmentation across our relatively small

study region, we attempted to assign latitude and longitude coordinates to each individual based on his or her village of origin. These assignments were made with the help of the online mapping resource, Global Gazetteer Version 2.1, found at <http://www.fallingrain.com/world/PP/>. In some cases an individual's village could not be found in these sources, although a village known to be nearby was present. In these cases the individual was assigned the geographic coordinates of the nearby village. In this way, 990 individuals could be assigned latitude and longitude values with confidence.

Recently, the development of remote sensing technologies has provided an opportunity for more direct measurement of UVR. Jablonski and Chaplin (2000) used data from the NASA Total Ozone Mapping Spectrometer (TOMS) to examine the relationship of UVR and skin reflectance directly. Although these researchers observed strong correlations with latitude and skin reflectance measured at various wavelengths, they also observed stronger correlations between UVR (measured as the minimum erythema dose, or UVMED) and skin reflectance at certain wavelengths when examining data from the northern and southern hemispheres. The advantage of using information such as that found in the NASA TOMS dataset is that it provides a direct measure of UVR. Although latitude and UVR are highly correlated, variation in cloud cover, altitude, or humidity may confound this relationship in localized regions. The effects of using latitude rather than a direct measure of UVR in this study will be discussed below.

Statistical Analysis

Mean and standard deviations for hair and skin M index values were calculated for males and females, linguistic phyla, islands, and neighborhoods. Standard two-sample t-tests were used to compare pigmentation between males and females as well as between linguistic phyla. In the case of unequal variances Satherwaite's corrected t-statistic was used. To control for sex-based differences in skin and hair pigmentation skin and hair M index values were z-standardized for sex. With the exception of actual male-female comparisons all statistical analyses were carried out using these standardized means, unless otherwise noted.

Differences between the group means of islands and neighborhoods were assessed using standard one-way analysis of variance. ANOVA was also used to compare neighborhood means at the island and phylum levels. Pairwise comparisons of neighborhoods within islands and phyla were carried out as t-tests with a Bonferroni correction for multiple comparisons. Interaction effects between the categorical and continuous variables in this study were also tested. The relationships between age, latitude, and longitude and pigmentation were assessed using a standard linear regression model. All statistical analyses were performed using the SAS 9.1 software package.

Results

Sex-based Differences

The results of standard two-sample t-tests comparing skin and hair pigmentation between males and females reveal that males are very highly significantly darker than females (male skin M index = 74.0, female skin M index = 71.2, $p < 0.0001$). Males were also darker than females

in hair pigmentation (male hair M index = 155.4, female hair M index = 151.2), although the value did not attain statistical significance ($p < 0.0537$). The significant differences between males and females in skin pigmentation and the near significant differences between males and females in hair pigmentation prompted the standardization of skin and hair M index values for sex, to avoid sex-related confounding of results. We also examined skin and hair pigmentation differences between males and females within phyla and islands. These results will be discussed in the appropriate sections below.

Effects of Increasing Adult Age

After standardizing for sex differences, mean hair M index was shown to decrease very slightly, but significantly ($R^2 = 0.0541$, $p < 0.0001$; 2.5% of total variation) with increasing adult age, while skin pigmentation increased fractionally, but significantly ($R^2 = 0.0281$, $p < 0.0001$). We also examined the relationship between age and pigmentation separately for males and females. Males showed a slightly stronger negative correlation between hair pigmentation and age after adulthood ($R^2 = 0.1139$, $p < 0.0001$) than females ($R^2 = 0.0106$, $p < 0.05$). When looking at skin pigmentation, however, females show a stronger correlation with age ($R^2 = 0.0873$, $p < 0.0001$) than males ($R^2 = 0.0060$, $p = \text{n.s.}$). In all cases, however, these age effects are weak, and although statistically significant may be of only marginal biological significance

Linguistic Phylum

There was no significant difference between AN and P speakers in terms of skin M index, although the two groups were highly significantly different for hair M index ($t = 3.81$ $p < 0.01$). Comparisons of males and females within each phylum revealed significant differences in skin M between AN males and females (female skin M index = 70.1, male skin M index = 75.1 $p < 0.0001$). No significant differences between Papuan males and females were observed for either skin or hair M index. Mean M index values for males and females within each phylum are shown in Table 2.2 .

Table 2.2: Mean skin and hair M index values for males and females in each linguistic phylum

Phylum	sex	n	skin M index		n	hair M index	
			mean	S.D.		mean	S.D.
Austronesian	F	331	70.1*	9.7	331	154.0	17.9
	M	374	75.1*	10.8	374	156.2	10.8
Papuan	F	94	74.1	10.6	94	147.8	21.1
	M	184	74.5	11.3	184	151.4	19

*significant male/female difference at $p < 0.0001$

Pigmentation of AN and P speakers was compared within each island. AN and P neighborhoods differed significantly from each other in hair pigmentation on both Bougainville ($t = 6.40$ $p < 0.0001$) and New Ireland ($t = 2.39$, $p < 0.05$). AN and P speakers were significantly different for skin pigmentation on New Ireland alone ($t = 2.93$, $p < 0.01$).

In addition to testing for differences between AN and P groups, we also tested to see if there was homogeneity in skin and hair pigmentation within the AN and P linguistic phyla. To do this, we compared skin and hair M index values among the 19 Austronesian-speaking neighborhoods and the 9 Papuan-speaking neighborhoods

(Table 2.3). Significant differences were found among the Austronesian neighborhoods for both skin and hair ($F = 4.40$, $df = 18$, $p < 0.0001$). The Papuan neighborhoods showed similar heterogeneity (skin: $F = 35.74$, $df = 8$, $p < 0.0001$; hair: $F = 7.91$, $df = 8$, $p < 0.0001$). However, much of this variation may be explained by variation between islands (see below). To address this, we used ANOVA to compare the mean pigmentation of neighborhoods within each phylum on an island-by-island basis. There was significant variation for both hair and skin pigmentation within New Britain's Austronesian and Papuan speaking neighborhoods. Significant skin pigmentation variation also existed between the Austronesian speaking groups on Bougainville. These results suggest that heterogeneity exists within AN and P speaking groups in the large islands of New Britain and Bougainville, but not for the smaller (and much narrower) island of New Ireland. This emphasizes the importance that island size and geographic complexity can have in maintaining and increasing genetic diversity. This inference is consistent with mtDNA and NRY findings in this same region (e.g. Merriwether, Friedlaender et al., 1999; Robledo et al., 2004; Scheinfeldt et al., 2004; Friedlaender et al., in press;).

Table 2.3: ANOVA results comparing islands, neighborhoods, neighborhoods within phyla and islands

Model	skin			hair		
	F	df	p	F	df	p
Island	222.23	5	0.0001	3.88	5	0.0017
Neighborhoods (entire sample)	34.96	27	0.0001	6.06	27	0.0001
Neighborhoods (within phyla)						
AN	37.30	18	0.0001	4.40	8	0.0001
P	35.50	8	0.0001	8.44	8	0.0001
Neighborhoods within islands						
New Britain	3.92	12	0.0001	8.74	12	0.0001
New Hanover	0.19	2	0.8383	1.00	2	0.3712
New Ireland	1.98	5	0.0852	1.79	5	0.1199
Bougainville	5.25	3	0.0022	18.81	3	0.0001
Neighborhoods within phyla within islands						
New Britain AN	4.28	8	0.0001	7.64	8	0.0001
New Britain P	4.06	3	0.0088	16.12	3	0.0001
New Hanover AN	0.19	2	0.8383	1.00	2	0.3712
New Ireland AN	1.53	2	0.2233	1.70	2	0.1910
New Ireland P	1.18	2	0.3126	0.40	2	0.6749
Bougainville AN	4.87	2	0.0111	2.46	2	0.0977

Island

Mean hair and skin M index values for each island, as well as for each major neighborhood within those islands, are shown in Table 2.4. Mean pigmentation for skin and hair was compared across the 6 major islands in our study using ANOVA (Table 2.3). Islands are significantly different for both hair M index ($F = 3.88$, $df = 5$, $p < 0.01$) and skin M index ($F = 222.23$, $df = 5$, $p < 0.0001$). Since the Bougainville skin pigmentation results are most distinctive from the other islands ($p < 0.0001$ for all pairwise comparisons), and might therefore be the sole reason for the significant inter-island difference we re-ran the analysis excluding Bougainville. Even after this

exclusion, the remaining islands are significantly different in skin pigmentation ($F = 68.47$, $df = 4$, $p < 0.0001$).

Mean skin and hair pigmentation values for males and females on each island are shown in Table 2.5 . Significant differences in skin M between males and females were observed on all islands but New Hanover. Significant differences in hair M index were only found on the island of New Britain.

In an effort to determine if pigmentation was homogeneous within islands, or if heterogeneity among the different neighborhoods on those islands existed, we also used ANOVA to compare mean pigmentation among neighborhoods on the islands of New Hanover, New Britain, New Ireland, and Bougainville, the four islands containing multiple neighborhoods (Table 2.3). We observed that there was significant variation in hair pigmentation on the islands of Bougainville ($F = 18.81$, $df = 3$, $p < 0.0001$) and New Britain ($F = 8.09$, $df = 13$, $p < 0.0001$). Significant variation in skin pigmentation was present on New Ireland ($F = 2.53$, $df = 5$, $p < 0.05$), New Britain ($F = 3.92$, $df = 12$, $p < 0.0001$), and Bougainville ($F = 5.25$, $df = 3$, $p < 0.01$).

Neighborhood

Neighborhoods are significantly different for hair ($F = 6.02$, $df = 27$, $p < 0.0001$) as well as for skin pigmentation ($F = 35.15$, $df = 27$, $p < 0.0001$). Again, thinking that the extremely dark-skinned Bougainvilleans might be influencing these results, we re-ran the analysis without the four Bougainville neighborhoods. After the exclusion of these neighborhoods, there was still significant evidence for differences in mean pigmentation

Table 2.4: Mean skin and hair M index values for islands¹ and neighborhoods²

Island	Neighborhood	n	skin M index		n	hair M index	
			mean (raw)	S.D. (raw)		mean (raw)	S.D. (raw)
PNG		21	-0.552 (67.9)	0.302 (9.4)	19	0.816 (160.1)	0.838 (15.3)
	Sepik	11	-0.418 (69.4)	0.967 (10.6)	10	0.869 (169.7)	0.679 (11.7)
New Britain		491	-0.513 (67.9)	0.041 (6.8)	476	0.625 (155.2)	0.960 (17.7)
	Arimegi Island (Kove)	45	-0.761 (65.0)	0.467 (5.3)	41	0.207 (157.5)	0.601 (11.0)
	Kariai (Anem)	25	-0.021 (72.2)	0.713 (7.5)	25	-0.175 (150.4)	0.473 (8.6)
	Pureling (Anem)	29	-0.627 (67.1)	0.681 (7.2)	25	-0.388 (147.5)	0.733 (13.2)
	Kisiluvi (Mamousi)	34	-0.274 (69.7)	0.569 (6.6)	32	0.103 (155.3)	0.886 (15.9)
	Lingite (Mamousi)	22	-0.388 (68.6)	0.593 (6.9)	22	0.363 (160.0)	0.873 (16.0)
	Welu (Mamousi)	11	-0.537 (67.3)	0.480 (5.2)	11	0.196 (157.5)	0.769 (13.2)
	other Mamousi	13	-0.530 (67.3)	0.704 (6.5)	13	0.363 (160.6)	0.971 (16.9)
	Loso (Nakanai)	17	-0.237 (70.3)	0.589 (5.5)	17	0.355 (160.2)	1.683 (29.7)
	Uasilau (Ata)	44	-0.547 (67.4)	0.543 (5.4)	44	-0.210 (150.2)	0.775 (14.2)
	Lugei (Ata)	20	-0.492 (67.6)	0.970 (9.7)	19	1.075 (173.3)	1.050 (18.2)
	Bileki (Nakanai)	96	-0.675 (65.7)	0.535 (6.2)	93	-0.507 (144.4)	0.665 (12.2)
	Ubili (Melamela)	36	-0.283 (69.3)	0.661 (6.6)	36	0.661 (165.0)	1.079 (20.2)
	Kuanua (Tolai)	18	-0.387 (68.3)	0.664 (6.3)	18	-0.076 (151.9)	1.08 (20.4)
Mussau		35	-0.723 (65.5)	0.684 (7.9)	33	0.448 (162.5)	0.544 (10.0)
	Kapugu	34	-0.728 (65.0)	0.662 (7.4)	32	0.495 (162.4)	0.563 (10.1)
New Hanover		102	0.433 (77.1)	0.705 (7.5)	98	0.018 (154.4)	0.856 (16.1)
	North Lavongai	73	0.501 (76.9)	0.685 (6.9)	69	0.109 (154.9)	0.854 (15.9)
	West Lavongai	10	0.379 (75.5)	0.776 (7.6)	10	0.251 (157.3)	0.968 (18.4)
	South Lavongai	13	0.408 (77.3)	0.889 (9.9)	13	-0.217 (150.3)	0.837 (14.2)
New Ireland		242	0.111 (74.2)	-0.152 (8.1)	222	0.724 (151.6)	1.264 (23.1)
	Tigak	27	-0.009 (72.6)	0.704 (8.4)	26	0.15 (156.4)	0.911 (15.9)
	Nailik	26	-0.061 (71.9)	0.573 (6.4)	20	-0.277 (148.9)	1.671 (29.8)
	Kabil (Kuot)	41	0.329 (76.2)	0.709 (7.8)	38	0.349 (147.4)	1.057 (19.1)
	Lamalaua (Kuot)	11	0.153 (75.0)	0.607 (6.2)	9	-0.755 (141.3)	1.72 (29.5)
	Notsi	21	0.286 (74.8)	0.852 (9.0)	20	0.435 (160.8)	1.002 (18.4)
	Madak	26	0.462 (78.0)	0.878 (8.9)	26	-0.205 (148.0)	1.447 (27.5)
		153	1.552 (89.8)	0.857 (9.5)	127	-0.205 (150.7)	0.809 (14.6)
Bougainville	Saposa Island (Saposa)	41	1.283 (86.1)	0.905 (9.9)	29	0.044 (154.3)	0.636 (11.4)
	Inivus (Teop)	10	2.052 (93.1)	0.807 (9.0)	9	0.616 (164.5)	0.660 (11.6)
	Sunahoara (Teop)	10	1.977 (94.6)	0.524 (6.5)	9	0.228 (158.5)	0.695 (12.0)
	Kukuavo (Aita)	32	1.901 (91.9)	0.734 (8.5)	30	-0.821 (138.6)	0.538 (8.9)

¹ Includes individuals belonging to the neighborhoods listed below each island heading, as well as individuals who could not be assigned to a particular neighborhood within the island

² Papuan-speaking neighborhoods are in boldface type

Table 2.5: Mean skin and hair M index values for males and females on each island

Island	sex	n	skin M index		n	hair M index	
			mean	S.D.		mean	S.D.
PNG	F	6	60.6*	6.9	5	154.9	19.3
	M	15	70.9*	8.8	14	162	13.9
New Britain	F	190	66.0***	6.9	186	152.5*	19.7
	M	301	69.2***	6.5	290	156.9*	16
New Hanover	F	61	76.1	6.5	60	153.4	16.9
	M	41	78.6	8.7	38	155.9	14.7
Mussau	F	17	61.3*	6.2	17	162.4	10.4
	M	18	69.5*	7.4	16	162.7	9.9
New Ireland	F	118	71.5***	7.4	111	152.6	23.5
	M	124	76.8***	7.8	111	150.6	22.7
Bougainville	F	62	86.6**	8.4	59	150.5	13.8
	M	91	92.1**	9.7	68	150.8	15.4

*significant male/female difference at $p < 0.05$

** significant male/female difference at $p < 0.001$

***significant male/female difference at $p < .0001$

among neighborhoods for both hair ($F = 5.46$, $df = 23$, $p < 0.0001$) and skin ($F = 13.99$, $df = 23$, $p < 0.0001$). Although this is likely influenced by between-island differences in pigmentation, the comparison of neighborhoods within islands (see above) demonstrates that these differences can be found even within islands.

Table 2.6 contains mean skin and hair pigmentation values for males and females in each of the 28 neighborhoods. In many cases, the samples sizes are too small to conduct t-tests of between sex differences. However, the general trend of darker males relative to females is still evident.

Table 2.6: Mean skin and hair M index values for males and females in each neighborhood

Island	Neighborhood	sex	n	skin M index		hair M index		
				mean	S.D.	n	mean	S.D.
PNG	Sepik	F	1	47.3	.	1	178.8	.
		M	10	71.6	8.1	9	168.7	11.9
New Britain	Arimegi Island (Kove)	F	15	61.5	6	14	156	15.2
		M	30	66.8	3.4	27	158.4	8.3
	Karaiai (Anem)	F	12	70.4	7.2	12	149	8.4
		M	13	73.8	7.8	13	151.6	8.8
	Pureling (Anem)	F	3	65.1	10.6	3	132.9	11.8
		M	26	67.3	6.9	22	149.5	12.3
	Kisiluvi (Mamousi)	F	16	65.9	5.3	15	152	13.9
		M	18	73.1	5.7	17	158.2	17.4
	Lingite (Mamousi)	F	10	64.1	6.2	10	155.6	18
		M	12	72.4	4.9	12	163.6	13.9
	Welu (Mamousi)	F	4	65.1	6.5	4	161.6	9.7
		M	7	68.5	4.3	7	155.1	15
	other Mamousi	F	3	73.2	10	3	161.3	11.3
		M	10	65.9	4.4	10	160.4	18.7
	Loso (Nakanai)	F	6	71.9	6.1	6	159.9	28.4
		M	11	69.4	5.3	11	160.4	31.8
	Bileki (Nakanai)	F	42	61.4	5.5	42	141.3	12.8
		M	54	69	4.5	51	147	11.3
	Lugei (Ata)	F	7	70.7	6	6	183.6	5.1
		M	13	66	10.6	13	168.5	20.2
Uasilau (Ata)	F	12	68.1	5.5	12	143	17.4	
	M	32	67.1	5.3	32	152.9	12.1	
Ubili (Melamela)	F	22	69	6.9	22	161.3	24.4	
	M	14	69.2	6.3	14	171	19.9	
Kuanua	F	9	69.9	5.6	9	144.9	26.2	
	M	9	66.7	6.8	9	158.8	9.6	
Mussau	Kapugu	F	17	61.3	6.2	17	162.4	10.4
		M	17	68.6	6.8	15	162.4	10.1
New Hanover	North Lavongai	F	48	76.6	6.5	47	152.6	17.2
		M	25	77.5	7.8	22	159.8	11.8
	West Lavongai	F	7	75.5	8.5	7	153	20.6
New Ireland	South Lavongai	M	3	75.5	6.7	3	167.4	4.7
		F	4	71.8	3.6	4	155.8	9.3
	Tigak	M	9	79.7	11	9	147.8	15.8
Bougainville	Tigak	F	12	65.7	5.8	12	158.1	11.5
		M	15	78.1	5.6	14	154.9	19.2
	Nailik	F	12	69.2	4.3	9	153.6	32.5
		M	14	74.2	7.2	11	144.9	28.4
	Kabil (Kuot)	F	16	72.5	6.6	16	146.1	21.8
		M	25	78.5	7.8	22	148.4	17.3
	Lamalaua (Kuot)	F	2	75.3	2.8	2	163.8	8.8
		M	9	74.9	6.9	7	134.9	30.4
	Notsi	F	14	72.1	8.7	13	154.5	19.5
		M	7	80.4	7.4	7	172.4	8.5
Madak	F	12	77.8	8.3	12	150.4	27.9	
	M	14	78.2	9.6	14	146	28.1	
Bougainville	Saposa Island (Saposa)	F	15	82.2	7.5	14	153.6	11.7
		M	26	88.4	10.5	15	155	11.6
	Inivus (Teop)	F	6	89.8	9.6	5	164.8	8.4
		M	4	98	6	4	164.2	16.2
	Sunahoara (Teop)	F	1	79.1	.	1	161.4	.
		M	9	96.3	3.8	8	158.1	12.7
Kukuavo (Aita)	F	17	87.6	8	17	140.1	5.9	
	M	15	96.7	6.1	13	136.5	11.8	

Latitude and Longitude

Latitude was not statistically related to hair ($R^2 = 0.0003$, $p < 0.5937$) or skin ($R^2 = 0.0000$, $p < 0.9805$) pigmentation, nor is there any suggestion of any such underlying tendency. The lack of a strong correlation between latitude and skin pigmentation here should not be seen as a contradiction of the well documented associations found across broader regions. In this sample, we have a very small range in tropical latitude (9.3°), and are using this as a proxy for UVR to test for its association with very large differences in pigmentation. If we restrict our analysis to only those samples from Island Melanesia proper (excluding PNG), we are looking at an even smaller range (1 to 7° latitude). One cannot expect very long term differential selection effects to result in well defined clines in small regions with histories of small population sizes, a series of in-migrations, combined with considerable isolation.

As a further illustration of this apparent contradiction, longitude does show some statistical association with pigmentation across our sample. Although the relationship between longitude and hair pigmentation was statistically significant ($p < 0.05$), it is extremely weak ($R^2 = 0.0057$). However, the relationship between skin pigmentation and longitude was quite strong ($R^2 = 0.3511$, $p < 0.0001$), with pigmentation increasing from west to east across the study region. As Bougainville has the highest mean skin pigmentation of any of the islands surveyed, and also lies at the extreme east of our study region, we suspected that it might explain this strong west-east cline in pigmentation values. When the analysis was re-run without the Bougainvilleans, we found that the correlation between longitude and pigmentation decreases dramatically, although it retains statistical significance ($R^2 = 0.0279$, $p < 0.0001$).

Interaction Effects

We also tested for interaction effects between our categorical variables of phylum, island, and neighborhood and the continuous variables age, latitude, and longitude. In these models we tested both the main effect of the continuous and categorical variables in question, but also for an interaction effect between the two. At the phylum level, significant interactions between phylum and longitude were observed for both skin ($F = 26.4$, $df = 1$, $p < 0.0001$) and hair ($F = 12.59$, $df = 1$, $p < 0.001$) pigmentation. Significant interactions between island and latitude ($F = 3.35$, $df = 5$, $p < 0.01$) and island and longitude ($F = 3.13$, $df = 5$, $p < 0.01$) were observed for skin pigmentation. There were also significant interactions between island and age ($F = 5.84$, $df = 5$, $p < 0.0001$) and island and longitude ($F = 2.58$, $df = 5$, $p < 0.05$) for hair pigmentation. Hair pigmentation was also influenced by interaction between neighborhood and age ($F = 2.55$, $df = 27$, $p < 0.0001$) and neighborhood and longitude ($F = 1.64$, $df = 17$, $p < 0.05$).

Discussion

Sex-based Differences

When we compared all of the males and females in this study (irrespective of phylum, island, or neighborhood affiliations) we observed very highly significant sexual dimorphism with respect to skin pigmentation, with males being darker than females ($p < 0.0001$). These differences persist when we compare males and females within the same island for five out of six islands. Although sample sizes were often too small to test for significant differences in males and females at the neighborhood level, the general trend

of males being darker is apparent. Males also tended to be darker than females for hair pigmentation, although this dimorphism is only of suggestive significance ($p < 0.0537$). The differences that we report in skin M index between males and females are consistent with a number of reports from the literature describing females as significantly lighter than males in most populations (e.g. Barnicot, 1958; Tobias, 1961; Conway and Baker, 1972; Byard and Lees, 1982; Harvey, 1985; and reviewed in van den Berghe and Frost, 1986; Frost, 1988). In their review of skin reflectance data, Jablonski and Chaplin (2000) reported a similar pattern. However, Wagner et al. (2002) did not find any such differences in their study of European Americans, Hispanics, or East Asians (although the Hispanic and East Asian sample sizes were small).

The reports of darker males relative to females in a number of studies have been explained as being due to hormonal differences between males and females or as the result of sexual selection favoring lighter females (Frost, 1984; Frost, 1994). Van den Berghe and Frost (1986) suggest that the male/female differences observed in many pigmentation studies are tied to a relationship between the reported decrease of female pigmentation around the time of menarche (as referenced in Robbins, 1991). Since skin pigmentation can increase with age, lighter skin may be seen as a sign of fecundity, causing males to view females with relatively lighter skin as desirable (van den Berghe and Frost, 1986). However, there are little data to confirm or refute the hypothesis that males actually choose their mates based on slight pigmentation differences that are only measurable by instrumentation and not obvious to the human eye.

It is also possible that any differences seen between the sexes are purely the result of different behavioral patterns that affect UVR exposure levels. The observations of

darker males in many studies are consistent with the idea that men may experience a greater lifetime UVR exposure, perhaps due to differences in clothing or activity patterns. For example, we observed that the large amount of time spent by men fishing on the ocean in outrigger canoes (where UVR exposure, due to reflectance off of the water, would be substantially higher than on land) could dramatically inflate between-sex differences in pigmentation. The result would be an increase in the facultative pigmentation of males relative to females, while constitutive pigmentation should remain the same. While our measurement site was chosen for its relatively low UVR exposure, some researchers suggest that the buttocks may be a more ideal measurement location to avoid seasonal variation in pigmentation that may occur due to facultative tanning (Lock-Andersen and Wulf, 1997).

Age

Hair pigmentation decreased slightly but significantly with age. Individuals noted to have white or graying hair were excluded from this analysis, so the negative relationship between age and hair M index cannot be easily explained by the presence of individuals with severely or obviously depigmented hair in the upper age categories. This suggests that melanin content of the hair does decrease with age independently of subjectively assessed graying. This may in part be due to the decline with age in the number of melanocytes present in the hair follicle, although this decline is typically associated with an increase in graying (Robbins, 1991). It is also possible that some individuals with a small number of gray hairs were included in the analysis. If a gray hair were included in the region measured by the DermaSpectrometer, this

could also explain the observed decrease in hair M index with age. This decrease in hair M index was more pronounced among males than among females.

We observed a slight increase in skin pigmentation with age in females ($R^2 = 0.0873$, $p < 0.0001$). Males showed no significant relationship. While many reports in the literature describe changes in pigmentation from birth through adolescence or young adulthood (e.g. Kalla and Tiwari, 1970; Conway and Baker, 1972; Williams-Blangero and Blangero, 1991), some have also examined changes in skin pigmentation throughout adulthood. Many of these suggest either no significant relationship between age and skin pigmentation (Harvey and Lord, 1978; Harvey, 1985), or a slight decrease in pigmentation with age (e.g. Conway and Baker, 1972; Frischano, Wainwright and Way, 1981). Again, although the inner arm was selected as a measurement site in this study because it is relatively protected from UVR (and was a culturally accepted measurement site, unlike the buttocks), it is still possible that over the course of many years facultative tanning could result in an overall increase in skin M index at this site. This could result in the weak, but positive, correlation between pigmentation and age observed here.

There are two additional factors that could explain these results, however. First, our sample is not equally distributed across all ages: over 80% of the individuals in our sample are below the age of 50. When the analysis is repeated including only individuals younger than 50 years of age, the correlation between skin pigmentation and age actually increases ($R^2 = 0.0308$, $p < 0.0001$), although the relationship still remains quite weak. Secondly, as we have seen that both skin and hair pigmentation differs significantly among islands, it is possible that the results of our age analysis could also be the result of significant differences in mean age between the islands. The mean age of our New

Britain sample (32.8 years) is significantly younger than that of all islands except Papua New Guinea ($p < 0.05$, after Bonferroni correction for multiple tests). This may also explain the significant interaction effect between age and island, and age and neighborhood (since neighborhoods are nested within islands) on hair pigmentation. Interestingly, Williams-Blangero and Blangero (1991) found that age-related changes in skin color varied among populations in eastern Nepal, although this study dealt only with age changes in skin pigmentation in individuals from age 3 to 21 years.

Linguistic Phylum

Constitutive pigmentation is under strong genetic regulation, and as such it was thought that observed differences between Austronesian and Papuan speakers may reflect pigmentation differences between the progenitors of these two groups, assuming there were, in fact, only two progenitor groups. However, studies of genetic variation between these two groups have led to conflicting results. For example although Giles et al. (1965) reported clear distinctions between Austronesian and Papuan speakers for Gm antigens in the Markham Valley of New Guinea such stark contrasts between the two groups were not found in Bougainville (Friedlaender and Steinberg, 1970) or in the coastal region of Madang in New Guinea (cf. Serjeantson et al., 1992).

Recent work suggests that the different conclusions about Austronesian/Papuan genetic distinctions may be due in part to different levels of admixture between Austronesian and Papuan speakers in the past. For example, the sharp differences in Gm antigen frequency observed by Giles et al. may be attributed to a recent arrival of

Austronesians to the region, and hence to less time for gene flow to reduce any pre-existing differences between the two groups (Friedlaender et al., in press). Studies where genetic differences between the two groups were observed to be less distinct may reflect greater amounts of admixture between the two groups. The adoption of a neighboring group's language may have also helped to blur biological distinctions between these groups, as in the Madak case. Another questionable factor in these comparisons is the presumed original homogeneity of Austronesians and (especially) the Papuans.

We did not observe significant differences between Austronesian and Papuan speakers in skin pigmentation, although the two groups were significantly different for hair pigmentation. Comparisons of AN and P speakers within the three islands for which we had samples of both groups (New Britain, New Ireland, and Bougainville) revealed significant differences in hair pigmentation on Bougainville and New Ireland. New Ireland was the only island to show significant differences between the two groups in skin pigmentation.

Our analyses demonstrated there were significant differences in skin and hair pigmentation between the AN and P neighborhoods in this study. After controlling for inter-island differences in pigmentation, we found evidence for heterogeneity in hair and skin pigmentation among the AN and P speaking neighborhoods of New Britain and in skin pigmentation among the AN speaking neighborhoods of Bougainville.

The fact that no significant within-island variation in skin pigmentation was found between Austronesian and Papuan speakers in Bougainville and New Britain suggests that perhaps sufficient admixture has occurred between the two groups to blur any

original skin pigmentation differences that may have existed. The significant difference in skin pigmentation between the two groups on New Ireland is especially interesting given the debated status of the Madak language as belonging to either the AN or P phylum. Ross (1994) believes that the Madak language, while traditionally being classified as AN, has an underlying structure that suggests that it was at one time a Papuan language. Over a period of time these proto-Madak speakers may have adopted features of the AN languages of their neighbors on New Ireland. Such conversions are also expected to be accompanied by gene flow between the two groups (Dutton, 1995). When the analysis is performed with Madak as an AN language, no significant differences between AN and P speakers are found in NI for either skin or hair pigmentation. As a group, Madak speakers represent the darkest people on New Ireland (M index = 78.0), and they also have the second-lowest M index values for hair (M index = 148.0) (the two Papuan speaking Kuot groups from Kabil and Lamalaua in our sample are the lightest with a mean hair M index of 146.2). The fact that Madak speakers are more similar to the one clear Papuan speaking group on New Ireland than they are to AN speakers suggests that they may have a more recent shared ancestry with P speaking groups such as the Kuot. It also suggests that the language conversion occurred fairly recently. It will be of great interest to compare the genetic variation of the Madaks to other AN and P groups in New Ireland.

Bougainville is the only other island on which significant AN/P pigmentation differences were detected. In this instance, only hair pigmentation differed between the groups. The variation in hair M index on Bougainville can be largely attributed to the lighter hair of the Papuan speaking Aita from Kukuavo (M index = 138.1) relative to the

darker Austronesian speaking Saposa (M index = 153.7) and Teop speakers from Inivus and Sunahoara (pooled M index = 162.0). Pairwise comparisons reveal that the Kukuavo are significantly different from all other Bougainvillean neighborhoods in hair pigmentation ($p < 0.0001$). The Aita until recently have been a relatively isolated group in the northern mountains of Bougainville, and are notable for lacking the mtDNA 9-bp deletion (often associated with the “Polynesian motif”) that characterizes all Austronesian populations (Merriwether and Friedlaender et al., 1999), as well as some Papuan-speaking groups (including those of southern Bougainville).

We observed significant interaction effects for both skin and hair pigmentation between the variables phylum and longitude. As we have already noted, Bougainville is unique in that it lies at the extreme east of our distribution and also has the darkest mean skin pigmentation. In addition, the Papuan-speaking Aita of Bougainville are notable for their very light hair. When Bougainvilleans are excluded, there is no significant interaction effect between phylum and longitude for either skin or hair pigmentation.

Island and Neighborhood Variation

Skin and hair pigmentation varies significantly at both the island and the neighborhood levels. Figure 2.2 shows skin M index variation for 27 of the 28 neighborhoods studied (the Sepik of New Guinea were excluded from the map for clarity). The five islands shown can be clearly distinguished by skin M index, despite some internal variation within these islands. The island with the lightest-skinned inhabitants is Mussau, located to the north of New Hanover and near the equator at

latitude – 1.4°. Individuals from New Guinea (not shown; mean skin M index = 67.9) and New Britain are also relatively lightly pigmented. Skin pigmentation increases in New Hanover and New Ireland, and increases again in Bougainville. The Bougainvilleans (mean skin M index = 89.8) are so dark compared to other groups in this study (mean skin M index = 70.6) that they were removed from the island-level analysis to determine if skin pigmentation still varied significantly among the remaining regions. Despite this, pigmentation differences among islands remained highly significant.

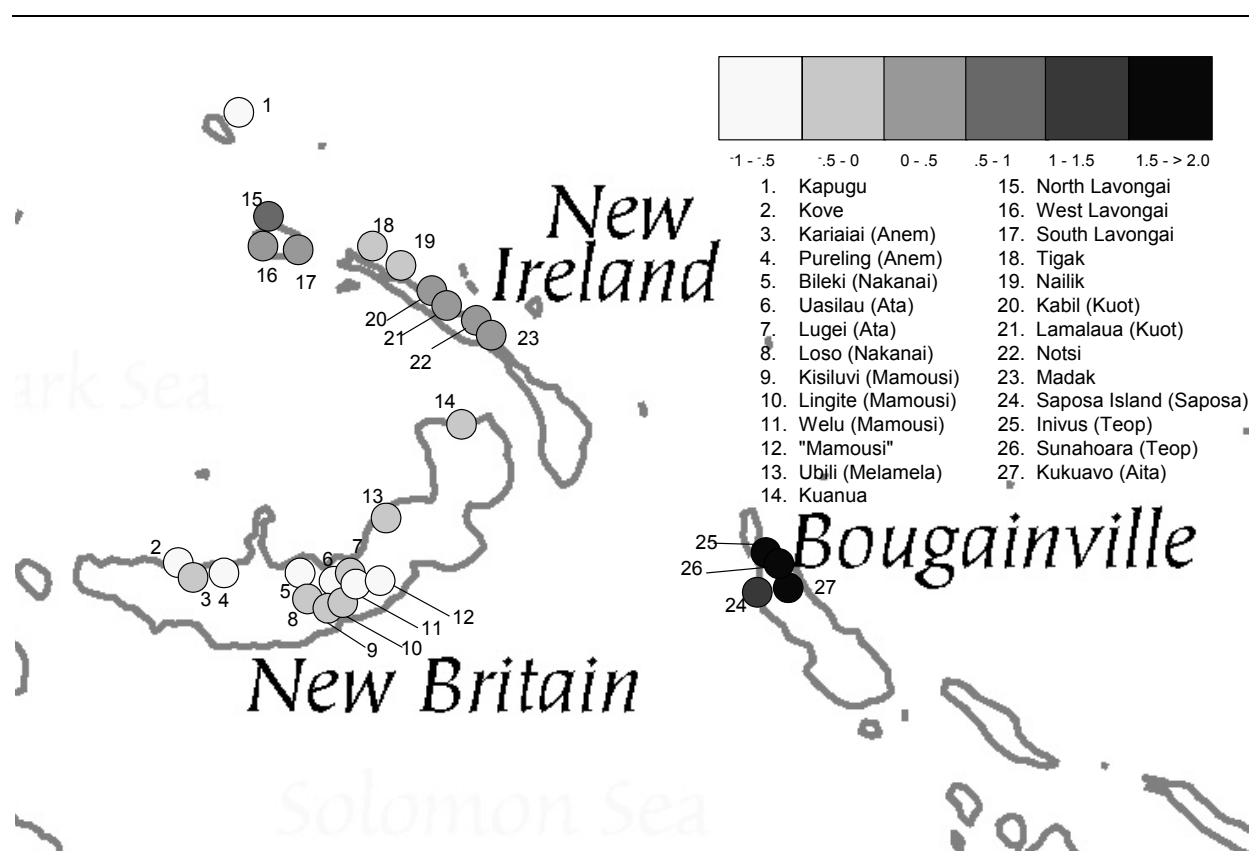


Figure 2.2: Skin M index for 27 of the 28 neighborhoods reported on in this study (the Sepik of New Guinea are not included). Intensity of the circle marking each population on the map corresponds to pigmentation as measured by the M index.

There is also significant variation in skin pigmentation among the 28 major neighborhoods was also observed. Neighborhoods from Bougainville were notably darker than populations from all other islands. When these Bougainvillean groups are excluded from the neighborhood-level analysis, neighborhoods still vary significantly with respect to both skin and hair M index. The heterogeneity within islands is consistent with the significant interaction effects observed for island and latitude and island and longitude. As such, it is not only *which* island an individual is from, but *where* on that island.

Although there was significant variation in skin pigmentation among the islands studied, significant within-island variation among neighborhoods was also observed for Bougainville, New Britain and New Ireland. This suggests that while regional distinctions in pigmentation may be important, homogeneity within those regions is not necessarily the norm. Studies such as Robledo et al. (2004) and Merriwether and Friedlaender et al. (1999) suggest that genetic variation within and between large islands follows a similar pattern.

Despite the large amount of variation in skin pigmentation in Island Melanesia, the inhabitants are darkly pigmented relative to other populations (See Figure 2.3 such as European Americans (mean skin M index = 29.5) and East Asians (mean skin M index = 32.6) and African Americans (mean skin M index = 53.4) and African Caribbeans (mean skin M index = 57.8) (Wagner et al., 2002; Shriver et al., 2003). This dark pigmentation is consistent with two of the many selection-based hypotheses regarding variation in human skin pigmentation. The first of these, the photoprotection hypothesis, is based on studies of the protective properties of melanin against UVR-induced damage such as sunburn and skin cancer (e.g. Pathak and Fitzpatrick,

1974). Specifically, protection provided by highly melanized skin would eliminate or minimize the painful effects of sunburn as well as reduce the risk of skin cancer.

Skin pigmentation in Bougainvilleans, African-Caribbeans, and African Americans

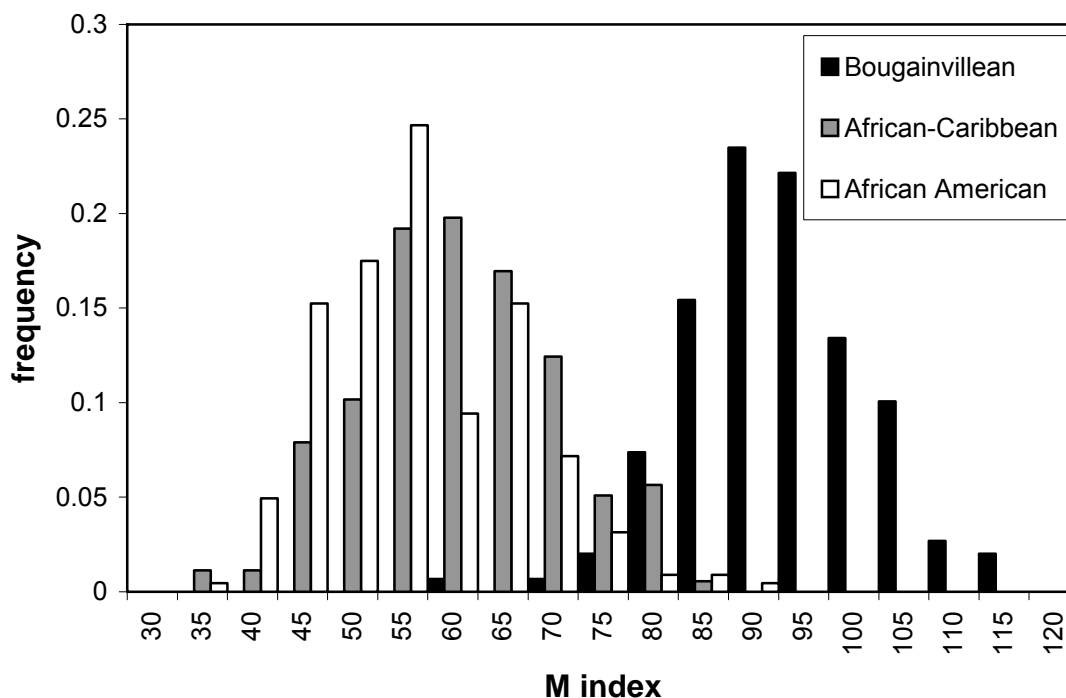


Figure 2.3: Distributions of skin M index values for the Bougainvilleans measured in this study, African Americans and African Caribbeans measured by Shriver et al. (2003).

The second major natural selection based hypothesis regarding normal variation in human skin pigmentation is also dependent on the photoprotective effects of melanin. In 1965, Hibbard and Smithells reported an association between fetal abnormalities and folate deficiencies. Later work sought to demonstrate a link between folate deficiencies and neural tube birth defects, as well as to demonstrate the effectiveness of folic acid supplementation at decreasing risk of such defects (e.g. Smithells et al. 1980; Laurence et al. 1981; Bower and Stanley, 1989; MRC Vitamin Study Research Group, 1991; Czeizel and Dudas, 1992; *inter*

alia). The folic acid hypothesis (Branda and Eaton, 1978) suggests that there is strong selection pressure to maintain a darkly pigmented skin in tropical regions to avoid the breakdown of folic acid and other metabolites via UVR exposure and hence an increase in neural tube birth defects (Jablonski, 1992) as well as nondisjunction errors during spermatogenesis (Mathur et al., 1977). Recently Flemming and Copp (1998) identified a link between folate deficiency and neural tube birth defects in the mouse, creating renewed interest in this hypothesis (e.g. Jablonski and Chaplin, 2000).

A third major hypothesis dealing with the evolution of human skin pigmentation variation is the Vitamin D Hypothesis (Loomis, 1967). However, as this hypothesis focuses mainly on the effects of positive selection for low melanin content in populations of the higher latitudes it is not pertinent to the populations under study here. It can be noted however, that the predictions of the vitamin D hypothesis are not inconsistent with darkly pigmented skin in regions of high UVR.

While there are other hypotheses regarding human pigmentation variation, these are the three that have received recent attention (Jablonski and Chaplin, 2000; Relethford, 1997). It should also be noted that these hypotheses are not necessarily mutually exclusive—it may be that all three together have helped to shape the global distribution of skin pigmentation. The suggestion of both Darwin (1871) and Diamond (1988) that sexual selection may also play a role in shaping human pigmentation variation should also not be ignored, although this idea is not testable using these data.

As predicted by the photoprotection and folic acid hypotheses, inhabitants of Island Melanesia do indeed have darkly pigmented skin, although there is a substantial amount of variation. It is possible that with respect to these two hypotheses there is a certain “melanin threshold” of skin pigmentation that represents a level of adaptive

pigmentation in a high UVR region. Individuals below that threshold are not well protected from UVR-induced photodamage to the skin or folic photolysis. Once that threshold is crossed selection ceases to be a strong force in constraining variation in skin pigmentation. Under this scenario, pigmentation above a certain level would not be a target of selection, while pigmentation below that level could be related to fitness. This may help to explain why, although in general darkly pigmented, the Island Melanesians show such extreme variation. A related concept was proposed by Chaplin (2004). Chaplin suggests that there may be a point at which it is simply not possible for pigmentation to increase. As humans near this melanin maximum, adaptation (in terms of melanin increase) should become increasingly slower. Rather than saying that Island Melanesians are bumping against this upper melanin boundary, we argue that as long as pigmentation is maintained above a certain protective level, individuals may vary in pigmentation with no (or minimal) negative effects on fitness. Certainly it is possible that under this scenario some groups may approach Chaplin's maximum—Bougainvilleans are a possible example.

An alternative is that the variation within the region that we observe is actually directly related to variation in UVR levels or to a population's duration of habitation in a particular UVR environment. Under this scenario, this variation may represent extremely localized adaptations to very specific environmental conditions. While we strongly believe that natural selection has shaped mean pigmentation levels in the region, our results are most consistent with the "melanin threshold" model, and we propose that variation within the region has been more strongly influenced by the population histories of these groups.

The variation in skin pigmentation we observe is consistent with the population history of the region and patterns of isolation and migration over the past few millennia. Reproductive isolation of these populations from one another may have led to differentiation in pigmentation at both island and neighborhood levels. This process could be the result of random genetic drift, founder effect, and differential gene flow, as well as possible localized sexual selection acting on pigmentation phenotype. While it is unlikely that these groups have remained completely isolated from each other for extensive periods of their histories, even partial isolation may have aided in this differentiation process. Bougainville is particularly interesting from this perspective, because it differs notably from the other regions considered here, and is also the most geographically isolated of the groups.

Our time spent in this region provided some anecdotal evidence for the possibility of localized sexual selection acting in Island Melanesian populations. For example, neighbors of the Anêm speaking people of West New Britain believed that Anêm women were ugly, and specifically that they had darker skin than themselves. This difference in pigmentation (and hence physical attractiveness) was cited as an argument against intermarriages with the Anêm. The skin reflectance measurements confirm the pigmentation observation (mean skin M index of Anêm women = 69.4, mean skin M index of Kove women = 61.5), but whether or not pigmentation was the driving factor limiting intermarriage with the Anêm remains to be seen. Similarly, in Bougainville Aita women were described as “disgusting” and darker than their neighbors (no significant difference in skin M index was observed), resulting in a similar claimed avoidance of intermarriage. This is interesting given that Bougainvilleans in general express a

preference for individuals with “clear black” skin—suggesting that perhaps there are other factors behind this avoidance practice. The Aita live in an isolated mountainous region of northern Bougainville, and although the Anêm had recently moved down to the coast they were originally an inland, rather than shore-based group. It may be that the resistance of coastal groups to intermarriage with supposedly “darker” populations such as the Aita or Anêm is really reflective of resistance to intermarriage with inland populations that are generally considered less sophisticated. Whatever the reason, however, this resistance may help to strengthen the genetic differences that have been observed between coastal and inland groups.

While the photoprotective and folic acid hypotheses predict dark pigmentation in tropical regions it seems as if the dark skin color of Bougainvilleans could not have been strictly controlled by the effects of natural selection alone. If it were, we would expect other Melanesian groups to display pigmentation levels similar to Bougainvilleans. However, the Bougainvilleans as a group are strikingly darker than the other Island Melanesians, as well as darker than previously measured African Americans and African Caribbeans, as shown in Figure 2.3. It should be noted that these African Americans and African Caribbeans are admixed populations, and that it is possible that other African or African derived populations would be more darkly pigmented. Also, African populations should not be considered to be homogeneous for skin pigmentation. The mean skin M index of 89.8 is remarkable in any context. It is possible that the remarkable pigmentation observed in the Bougainvillean populations has also been shaped by either drift and/or sexual selection.

Hair M index variation across the populations is distinctly different from variation in skin pigmentation, though still highly variable. Figure 2.4 shows hair M index values plotted for 27 of the 28 major populations in this study (as before, the Sepik of New Guinea have been excluded, sex standardized $M = 169.7$). Although the islands differ significantly by hair M index, there is also substantial variation among the populations within each island, particularly for Bougainville and West New Britain. As with skin pigmentation, we also observed a significant interaction effect for island and longitude on hair pigmentation. We feel that this is consistent with the within island heterogeneity observed for hair pigmentation.

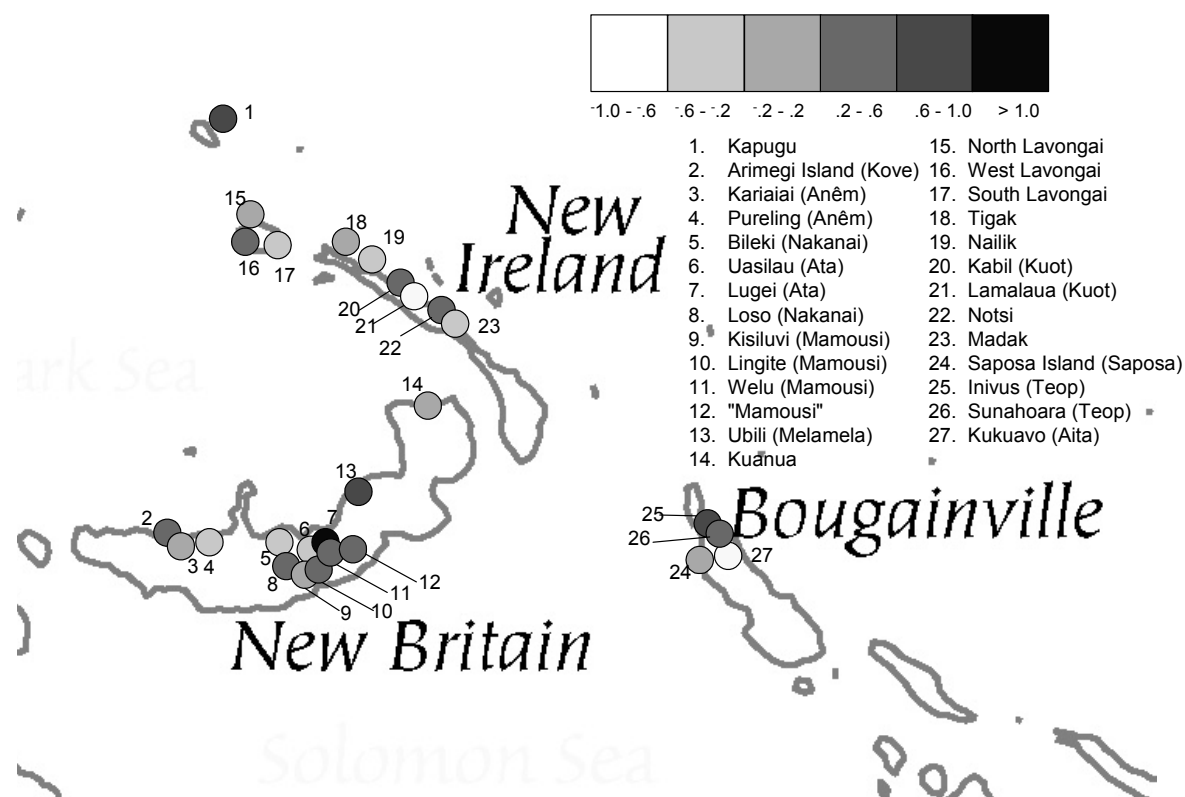


Figure 2.4: Hair M index for 27 of the 28 neighborhoods reported on in this study (the Sepik of New Guinea are not included). Intensity of the circle marking each population on the map corresponds to pigmentation as measured by the M index

An interesting hair phenotype that is sometimes seen in Island Melanesia (as well as among Australian Aborigines) is “blondism”, in which individuals exhibit the characteristic darkly pigmented skin of the region while also having blond hair. This trait was most commonly observed in children whose hair generally darkens around puberty (Robbins, 1991). However, in some cases blondism does persist into adulthood, although the hair does appear somewhat darker than what is seen in children. The strikingly light hair of the Aita (see Figure 4) is partially due to the high incidence of blondism among this group.

While there has been a large body of work devoted to explaining variation in skin pigmentation, especially via natural selection, much less work has been done to explain normal variation in hair pigmentation. It is possible that variation in hair pigmentation is mediated not so much by natural selection as it is by sexual selection. While no association between skin and hair M index was found in the combined sample ($R^2 = 0.0003$, $p < 0.5641$), it is interesting to consider how natural selection-based constraints on skin pigmentation in this tropical region may have indirectly limited variation in hair pigmentation.

At least some pigmentation candidate genes, such as the melanocortin 1 receptor (*MC1R*), can have an effect on both hair and skin pigmentation. Perhaps if those genes were under functional constraint to maintain a darkly melanized skin in this high UVR region then hair pigmentation may have been influenced secondarily. While mutations that led to slightly lower levels of melanin in hair pigmentation may not have been selected against in this region, mutations leading to decreased melanin levels in both hair and skin would have been. In this way, variation in hair pigmentation may be generally constrained by selection acting on skin pigmentation. The phenomenon of blondism may be due to a mutation that only affects hair color.

Latitude and Longitude

We found no correlation between either skin or hair pigmentation and latitude in this study, which is not surprising given that this sample covers such a small range of latitude (9.3°), and most studies stressing the relationship between latitude and

pigmentation have examined the relationship across a global scale (e.g. Relethford, 1997; Jablonski and Chaplin, 2000). Despite the high levels of variation in skin M index, the relatively dark skin pigmentation of individuals in this study is consistent with expectations for a tropical region where natural selection may have acted to maintain dark skin as a protection from UV-induced skin damage or folic acid photolysis.

It is important to emphasize that latitude may not be a suitable proxy for UVR in this region, due to effects of cloud cover, humidity, and varying elevations. Thus, the lack of correlation between latitude and skin pigmentation does not necessarily imply a lack of correlation between pigmentation and UVR. With that in mind, however, it is difficult to imagine natural selection exhibiting such fine-tuned regulation of pigmentation at these very localized levels. One useful test of this would be to determine if UVR levels on Bougainville are dramatically elevated relative to elsewhere in the region. If UVR is unusually high there, where pigmentation is darkest, then the case for localized selection of pigmentation phenotype is strengthened.

Jablonski and Chaplin (2000) and Chaplin (2004) have recently made use of UVR data collected as part of NASA's Total Ozone Mapping Spectrometer (TOMS) project using UVR data collected from the Nimbus-7 satellite (1978-1993). We examined a smaller TOMS dataset for the years 1997 to 2002 collected by the Earth Probe satellite (http://toms.gsfc.nasa.gov/ep_toms/ep_v8.html) to determine mean UVMED levels across Island Melanesia. UVMED is a measure of the minimum amount of UV exposure required to produce a reddening in the skin of a lightly-pigmented individual. While the TOMS dataset is the best that is publicly available, we should note that one potential problem with using these data in our study region is that TOMS data is reported as

UVMED in cells of 1° latitude by 1.25° longitude. We saw that UVMED varied across our region, but that contrary to predictions Bougainville did not show the highest mean UVMED value (see Table 2.7). Interestingly, the highest UVMED value was observed for the island of Mussau (mean UVMED = 292), which in our study showed the lowest mean skin pigmentation (mean skin M = 65.5). UVMED values for our two other lightly pigmented islands, PNG and New Britain, were lower than those observed for the islands of New Ireland and New Hanover, but comparable to those of Bougainville.

Table 2.7: Mean M and UVMED values by island

island	standardized skin M index	raw skin M index	UVMED
PNG	-0.552	67.9	275
New Britain	-0.513	67.9	274
Mussau	-0.723	65.5	292
New Hanover	0.433	77.1	289
New Ireland	0.111	74.2	288
Bougainville	1.552	89.8	274

Hair M index increased slightly but significantly from west to east, while skin M index showed a much stronger increase over the same distance. However, it should be noted that individuals sampled from Bougainville are at the extreme east of our longitudinal distribution and exhibit the darkest skin M index values, making it possible that the Bougainvilleans are responsible for much of this correlation between longitude and skin M index. When Bougainvilleans are excluded from the analysis, longitude is no longer a significant predictor of hair M index, and becomes a weak, although still a significant ($R^2 = 0.0225$, $p < 0.0001$) predictor of skin M index.

Conclusions

Along with a suite of many other genetically controlled phenotypes, pigmentation is a trait that shows remarkable variation and structure within Island Melanesia. Not only was there great variation among islands, but the larger, more rugged islands showed significant internal variation in both skin and hair pigmentation. The geographic patterning of this variation cuts across Austronesian/Papuan boundaries (the possible exception being hair pigmentation in Bougainville), reflecting the complex, but very important, relationship between language, geography, and biological appearance in Island Melanesia. In this region, the structure (or pattern) of pigmentation variation echoes the same pattern of variation being elucidated at other loci, suggesting the same overriding effect within this region of demographic determinants (migration, drift, and population history).

The extreme variation exhibited within this sample set is not a contradiction to global studies that have demonstrated a strong correlation between UVR and skin pigmentation, but it does highlight some unanswered questions and raise some new hypotheses. Because Northern Island Melanesia extends from only 1 to 7 degrees S latitude, it occupies a high UVR environment that should confer an extremely strong selective constraint on skin pigmentation, whatever selective model one favors. In the simplest deterministic model, one might expect everyone there to be as darkly pigmented as Bougainvilleans. That they are not so heavily pigmented, reflecting the strong effects of population dynamics and history, suggests the following alternatives. First, although people have lived in Northern Island Melanesia for the past 40,000 years, this may be too short an interval for them to have responded completely to the UVR related selection.

How fast UVR related selection has an effect is very poorly understood. A second alternative hypothesis concerning the variation in this region is that there is a threshold effect with regard to the intensity of UVR related selection. That is, above a certain “melanin threshold” the force of natural selection on skin pigmentation in high UVR regions may be relatively weak or nonlinear. This hypothesis has some resonance with earlier work relating to critical UVR levels for stimulation of vitamin D₃ synthesis. Such a selective threshold would allow for other factors, such as gene flow, genetic drift, and/or sexual selection to play a role in shaping localized pigmentation.

Acknowledgements

This work was supported in part by grants from The Wenner-Gren Foundation for Anthropological Research, The National Geographic Society Exploration Fund, the National Science Foundation, and Pennsylvania State University. We thank two anonymous reviewers for their comments on this paper. We also wish to express our gratitude to the people of Papua New Guinea, who were very interested participants in the project. Particular thanks go to Boston Tustikai, Noah Lurang, Jerry Nauta, Luke Angis, and Paul Kave.

Literature Cited

- Allen J. 2003. Discovering the Pleistocene in Island Melanesia. In: Sand C., editor. Pacific Archaeology: Assessments and Anniversary of the First Lapita Excavation (July 1952) Koné, Nouméa, 2002. Nouméa, New Caledonia: Le Cahiers de l'Archéologie en Nouvelle-Calédonie. p 33-42.
- Allen J. 1996. The pre-Austronesian settlement of Island Melanesia. In: Goodenough WH, editor. Prehistoric settlement of the Pacific. Philadelphia: American Philosophical Society. p 11-27.
- Barnicot N. 1958. Reflectometry of the skin in southern Nigerians and in some mulattoes. Hum Biol 30:150-160.
- Bower C, and Stanley F. 1989. Dietary folate as a risk factor for neural tube defects: evidence from a case-control study in Western Australia. M J Aust 150:613-619.
- Bowler JM, Johnston H, Olley JM, Prescott JR, Roberts RG, Shawcross W, and Spooner NA. 2003. New ages for human occupation and climatic change at Lake Mungo, Australia. Nature 421:837-840.
- Branda R, and Eaton J. 1978. Skin color and nutrient photolysis: an evolutionary hypothesis. Science 201:625-626.
- Byard P, and Lees F. 1982. Skin colorimetry in Belize. II. Inter- and Intra-population variation. Am J Phys Anthropol 58:215-219.
- <http://www.calle.com/world/PP/index.html>

- Chaplin G. 2004. Geographic distribution of environmental factors influencing human skin coloration. *Am J Phys Anthropol* 125:292-302.
- Conway D and Baker P. 1972. Skin reflectance of Quechua Indians: the effects of genetic admixture, sex, and age. *Am J Phys Anthropol* 36:367-282.
- Cowles R.B. 1959. Some ecological factors bearing on the origin and evolution of pigment in the human skin. *Am Nat* 93:283-293.
- Czeizel AE, and Dudas I. 1992. Prevention of the first occurrence of neural tube defects by periconceptual vitamin supplementation. *N Engl J Med* 327:1832-1835.
- Darwin C. 1871. *The Descent of Man, and Selection in Relation to Sex*. London: J. Murray.
- Diamond J. 1988. Survival of the Sexiest. *Discover* May 74-81.
- Diamond J. 1992. *The Third Chimpanzee*. New York: Harper Collins Publishers. p 120.
- Diffey B, Oliver R, and Farr P. 1984. A portable instrument for quantifying erythema induced by ultraviolet radiation. *Br J Dermatol* 111:663-72.
- Dutton T. 1995. Language contact and change in Melanesia. In: Bellwood P, Fox J, Tryon D editors. *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University. p 17-38.
- Flemming A, and Copp A. 1998. Embryonic folate metabolism and mouse neural tube defects. *Science* 280:2107-2109.
- Friedlaender J, Gentz F, Thompson F, Kaestle F, Schurr T, Koki G, Mgone C, McDonough J, Smith L, and Merriwether DA. In press. mtDNA haplogroup distributions in New Guinea and Island Melanesia. In: Pawley A, Attenborough

- R, Golson J, and Hide R, editors. *Papuan Pasts: studies in the cultural, linguistic, and biological history of the Papuan-speaking peoples*. Adelaide: Crawford House.
- Friedlaender J. 1975. *Patterns of Human Variation: The demography, genetics, and phenetics of Bougainville Islanders*. Cambridge: Harvard University Press.
- Friedlaender J. 1987. *Conclusions*. In: Friedlaender J. editor. *The Solomon Islands Project: A long-term study of health, human biology, and culture change*. Oxford: Clarendon Press.
- Friedlaender J and Steinberg G. 1970. Anthropological significance of gamma globulin (Gm and Inv) antigens in Bougainville Island, Melanesia. *Nature* 228:59-61.
- Frischano AR, Wainwright R, and Way A. 1981. Heritability and components of phenotypic expression in skin reflectance of Mestizos from the Peruvian lowlands. *Am J Phys Anthropol* 55:203-208.
- Frost P. 1994. Geographic distribution of human skin colour: a selective compromise between natural selection and sexual selection? *Human Evolution* 9:141-153.
- Frost P. 1988. Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspect Biol Med* 32:38-58.
- Giles E, Ogan E, and Steinberg A. 1965. Gamma-globulin factors (Gm and Inv) in New Guinea: anthropological significance. *Science* 150:1158-1160.

- Harvey R, and Lord J. 1978. Skin colour of the Ainu of Hidaka, Hokkaido, Northern Japan. *Ann Hum Biol* 5:459-467.
- Harvey R. 1985. Ecological factors in skin color variation among Papua New Guineans. *Am J Phys Anthropol* 66:407-416.
- Hibbard E, and Smithells R. 1965. Folic acid metabolism and human embryopathy. *Lancet* 1:1254.
- Jablonski N, and Chaplin G. 2000. The evolution of human skin coloration. *Human Evolution* 39:57-106.
- Kalla A, and Tiwari S. 1970. Sex differences in skin colour in man. *Acta Genet Med. Gemellol* 19:472-476.
- Koki G, and Friedlaender J. 2004. Marital migration rates, malarial epidemiology, and biological diversity within Island Melanesia [abstract]. In: Program of the 73rd Annual Meeting of the American Association of Physical Anthropologists; 2004 April 14-17; Tampa.
- Kollias N, Sayer R, Zeise L, and Chedekel M. 1991. Photoprotection by melanin. *J Photochem Photobiol B*. 9:135-160.
- Laurence KM, James N, Miller MH, Tennant GB, and Campbell H. 1981. Double-blind randomised controlled trial of folate treatment before conception to prevent recurrence of neural-tube defects. *Br Med J (Clin Res Ed.)* 282:1509-1511.

- Leavesley M, Bird M, Fifield L, Hausadaen P, Santos G, and di Tada M. 2002. Buang Merabak: Early evidence for human occupation in the Bismarck Archipelago, Papua New Guinea. *Australian Archaeology* 54:55-57.
- Leavesley M, and Allen J. 1998. Dates, disturbances and artefact distributions: Another analysis of Buang Merabak, a Pleistocene site on New Ireland, Papua New Guinea. *Archaeology in Oceania* 33:63-82.
- Lock-Andersen J, and Wulf H. 1997. Seasonal variation of skin pigmentation. *Acta Derm Venereol* 77:219-221.
- Loomis W. 1967. Skin pigment regulation of vitamin-D biosynthesis in man. *Science* 157:501-506.
- Mackintosh J. 2001. The antimicrobial properties of melanocytes, melanosomes, and melanin and the evolution of black skin. *J Theor Biol* 211:101-113.
- Mathur U, Datta S, and Mathur B. 1977. The effect of aminopterin-induced folic acid deficiency on spermatogenesis. *Fertil Steril* 28:1356-1360.
- Merriwether DA, Friedlaender J, Mediavilla J, Mgone C, Gentz F, and Ferrell R. 1999. Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melanesia. *Am J Phys Anthropol* 110:243-270.
- MRC Vitamin Study Research Group. 1991. Prevention of neural tube defects: Results of the Medical Research Council Vitamin Study. *Lancet* 338:131-137.

- Murray F. 1934. Pigmentation, sunlight and nutritional disease. *Am Anthropol* 36:438-445.
- Pathak M and Fitzpatrick T. 1974. The role of natural photoprotective agents in human skin. In: Pathak M, editor. *Sunlight and Man*, Tokyo: University of Tokyo Press. p 725-750.
- Pavrides C, and Gosden C. 1994. 35,000 year-old sites in the rain forests of West New Britain, Papua New Guinea. *Antiquity* 69:605-610.
- Post P, Daniels F, and Binford R. 1975. Cold injury and the evolution of "white" skin. *Hum Biol* 47:65-80.
- Relethford J. 1997. Hemispheric differences in human skin color. *Am J Phys Anthropol* 104:449-457.
- Robins A. 1991. *Biological perspectives on human pigmentation*. New York: Cambridge University Press.
- Roberts D, and Kahlon D. 1976. Environmental correlations of skin colour. *Ann Hum Biol* 3:11-22.
- Roberts R, Jones R, Spooner N, Head M, Murray A, and Smith M. 1994. The human colonization of Australia: optical dates of 53,000 and 60,000 years bracket human arrival at Deaf Adder Gorge, Northern Territory. *Quaternary Geochronology* 13:575-583.

- Robledo R, Scheinfeldt L, Merriwether DA, Thompson F, and Friedlaender J. 2004. Pattern of Genetic Diversity in Island Melanesia. *Hum Biol* 75:941-949.
- Ross M. 1994. Areal phonological features in north central New Ireland. In: Dutton T and Tryon D, editors. *Language Contact and Change in the Austronesian World*. New York: Mouton de Gruyter. p 551-572.
- Scheinfeldt L, Friedlaender J, Thompson F, Robledo R, Merriwether DA, Beggs W, Bender W, Karafet T, Hammer M, and Lorenz J. 2004. Y chromosome variation in Melanesian populations [abstract]. In: Program of the 73rd Annual Meeting of the American Association of Physical Anthropologists; 2004 April 14-17; Tampa.
- Serjeantson S, and Gao X. 1995. *Homo sapiens* is an evolving species: Origins of the Austronesians. In: Bellwood P, Fox J, Tryon D editors. *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University. p 17-38.
- Serjeantson S, and Board B. 1992. Population genetics in Papua New Guinea. In: Attenborough R, Alpers M, editors. *Human biology in Papua New Guinea: The small cosmos*. Oxford: Oxford Science Publications. p 198-233.
- Sheehan J, Cragg N, Chadwick, C, Potten C, and Young A. 2002. Repeated ultraviolet exposure affords the same protection against DNA photodamage and erythema in human skin types II and IV but is associated with faster DNA repair in skin type IV. *J Invest Dermatol* 118:825-829.

Shriver M , Parra E, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameraon N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R. 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet* 112:387-399.

Shriver M, and Parra E. 2000. Comparison of narrow-band reflectance spectroscopy and tristimulus colorimetry for measurements of skin and hair color in persons of different biological ancestry. *Am J Phys Anthropol* 112:17-27.

Smithells RW, Sheppard S, Schorah CJ, Seller MJ, Nevin NC, Harris R, Read AP, and Fielding DW. 1980. Possible prevention of neural-tube defects by periconceptional vitamin supplementation. *Lancet* 1:339-40.

Spriggs M. 1995. The Lapita Culture and Austronesian prehistory in Oceania. In: Bellwood P, Fox J, Tryon D editors. *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University.

Spriggs M. 1997. *The Island Melanesians*. Cambridge, MA: Blackwell Publishers.

Summerhayes G, and Allen J. 1993. The transport of Mopir obsidian to Late Pleistocene New Ireland. *Archaeology in Oceania* 28:145-149.

Tasa G, Murray C, and Boughton J. 1985. Reflectometer reports on human pigmentation. *Curr Anthropol* 26:511-512.

Tobias P. 1961. Studies on skin reflectance in Bushmen-European hybrids. *Proceedings, Second International Congress of Human Genetics, Rome*. 461-471.

http://toms.gsfc.nasa.gov/eptoms/ep_v8.html

Van den Berghe P, and Frost P. 1986. Skin color preference, sexual dimorphism and sexual selection: a case of gene culture co-evolution? *Ethnic and Racial Studies* 9:87-113.

Wagner J, Parra E, Norton H, Jovel C, and Shriver M. 2002. Skin responses to ultraviolet radiation: effects of constitutive skin pigmentation, sex, and ancestry. *Pigment Cell Res* 15:385-390.

Walter H. 1971. Remarks on the environmental adaptation of man. *Humangenetik* 13:85-97.

Wasserman H. 1965. Human pigmentation and environmental adaptation. *Archive Env Health* 11:691-694.

Williams-Blangero S, and Blangero J. 1991. Skin color variation in Eastern Nepal. *Am J Phys Anthropol* 85:281-291.

Wickler S, Spriggs M. 1988. Pleistocene human occupation of the Solomon Islands, Melanesia. *Antiquity* 62:703-706.

Wurm S. 1982. *Papuan Languages of Oceania*. Philadelphia: J Benjamins of North America.

Chapter 3

Pigmentation Candidate Gene Variation in Island Melanesia

Introduction

The distribution of human skin pigmentation variation around the world follows general geographic clines. The association between skin pigmentation and latitude, or more specifically, ultraviolet radiation (recently summarized by Relethford, 1997; Jablonski and Chaplin, 2000; Chaplin, 2004) has been the basis for a number of hypotheses suggesting that normal variation in human skin pigmentation has been shaped by natural selection (e.g. Cowles, 1959; Wasserman, 1965; Loomis, 1967; Post et al., 1975; Branda and Eaton, 1978; Mackintosh, 2001;), although sexual selection is also a possible explanation (Darwin, 1871; Diamond, 1988; Frost, 1988; Diamond, 1992). Most discussions of human skin pigmentation variation tend to focus on broad, intercontinental differences, but variation within smaller regions can also be observed. One such example comes from Island Melanesia, a region well-known for its genetic and linguistic diversity (Friedlaender, 1975; Wurm, 1982; Dutton, 1995; Merriwether et al., 1999; Robledo et al., 2004; Friedlaender et al., 2005). This variation in skin pigmentation in Island Melanesia was previously described in Chapter Two and in Norton et al. (in press). The dark pigmentation of Island Melanesians is consistent with many natural selection-based hypotheses, but there is also substantial variation within the region that does not correspond to variation in ultra-violet radiation (UVR).

This paper explores the relationship between polymorphisms in pigmentation candidate genes and the phenotypic variation observed in an Island Melanesian sample. Ten SNPs in six pigmentation candidate genes (*TYR*, *TYRP1*, *ASIP*, *OCA2*, *MATP*, and *MC1R*) were typed in 647 individuals from the original sample of 1135 described in Chapter Two and Norton et al. (in press). For comparison, these SNPs were also typed in 60 individuals each from five

geographically diverse populations (East Asians, Native Americans, South Asians, West Africans, and Europeans).

Materials and Methods

Samples

Methods of sample collection, pigmentation measurement, and classification into island, phylum, and neighborhood categories were carried out as described in Norton et al. Due to time and financial constraints only a subset of that original sample (647 of 1135 individuals) was genotyped. These individuals were selected to maximize the sample sizes from the three primary islands: New Britain, New Ireland, and Bougainville. The Austronesian and Papuan speaking neighborhoods that were genotyped are listed in Table 3.1. The raw skin and hair M index values for each island and neighborhood are shown in Table 3.2. As only small numbers of individuals from New Guinea and New Hanover were genotyped, they are excluded from island-level analyses.

Table 3.1: Sample composition of Austronesian and Papuan speakers

Austronesian	n	Papuan	n
Kapugu	32	Aita	35
Kove	45	Anem	52
Kuanua	15	Ata	51
Mamusi	51	Kuot	43
Melamela	36	Madak	26
Nailik	22	"Papuan"	12
Nakanai	30		
Notsi	17		
Saposa	40		
Teop	23		
Tigak	25		
"Austronesian"	50		
Total	386		219

Table 3.2: Raw skin and hair M index values for neighborhood and island samples. Papuan-speaking neighborhoods are shown in boldface

Island	Neighborhood	n	raw M	skin		hair	
				SD	raw M	SD	raw M
PNG		7	71.3	8.2	163.9	11.6	
New Britain		288	67.9	6.4	155.9	17.7	
	Arimegi Island (Kove)	45	65.0	5.3	157.5	11.0	
	Kariai (Anem)	24	72.4	7.7	150.2	8.7	
	Pureling (Anem)	27	66.5	7.1	147.8	13.5	
	Kisiluvi (Mamousi)	24	69.2	6.2	155.4	17.9	
	Lingite (Mamousi)	12	69.2	6.4	165.4	17.8	
	Welu (Mamousi)	10	68.4	3.6	157.9	13.8	
	Loso (Nakanai)	11	70.6	6.5	163.2	27.2	
	Uasilau (Ata)	39	67.4	5.3	150.8	14.9	
	Lugei (Ata)	12	67.0	5.3	78.0	8.9	
	Bileki (Nakanai)	19	66.0	6.8	142.7	15.1	
	Ubili (Melamela)	36	69.1	6.6	165.0	20.2	
	Kuanua (Tolai)	15	68.1	6.9	153.1	21.5	
	Mussau		32	65.8	6.6	163.2	9.8
Kapugu		32	65.8	6.6	163.2	9.8	
New Hanover		7	80.5	5.5	148.8	8.4	
New Ireland		158	74.4	8.5	151.3	22.9	
	Tigak	25	72.9	8.6	156.0	16.5	
	Nailik	21	72.8	6.6	145.4	30.8	
	Kabil (Kuot)	33	76.1	8.1	148.9	18.7	
	Lamalaua (Kuot)	10	74.8	6.5	142.5	31.3	
	Notsi	17	74.6	9.8	161.0	20.5	
	Madak	26	78.0	8.9	148.0	27.5	
					149.5	14.5	
Bougainville		135	89.6	9.0			
	Saposa Island (Saposa)	36	86.2	8.0	153.9	11.6	
	Inivus (Teop)	10	93.1	9.0	164.5	11.6	
	Sunahoara (Teop)	10	94.6	6.5	158.5	12.0	
	Kukuavo (Aita)	32	91.9	8.5	138.6	8.9	

For comparison, 60 individuals each from five geographically diverse populations were also genotyped for the panel of ten pigmentation candidate SNPs. This world panel was composed of the following samples: 62 individuals from Valencia, Spain to represent European populations; 67 Nahua individuals from Guerrero, Mexico to represent Native American populations; 10 (each) Chinese and Japanese samples from the Coriell Institute Human Variation Cell Repository and 46 unadmixed Chinese individuals from Trinidad and Tobago (collected by

Tamiko Brown, UWI) to represent East Asian populations; 22 Indians from Andhra Pradesh and 45 unadmixed Indians from Trinidad and Tobago (also collected by Tamiko Brown, UWI) to represent South Asian populations; and seventy Gullah African-American individuals from South Carolina estimated to have 100% West African ancestry by Parra et al. (2001) to represent West African populations.

Laboratory Methods

As the quantity of the Island Melanesian DNA available for use in this study was limited (25 μ l @ < 10ng/ μ l), purified DNA from each sample was subject to whole genome amplification (WGA) using the GenomiPhi DNA Amplification Kit from Amersham Biosciences. The GenomiPhi kit utilizes the isothermal properties of ϕ 29 polymerase and random hexamer primers to yield 10,000-50,000-fold amplification from small quantities of starting material in a method commonly referred to as multiple strand displacement amplification (Lizardi et al., 1998; Dean et al., 2001; Davis et al., 2002). Amplifications were carried out in a 20 μ l reaction according to the manufacturer's instructions. The results of the GenomiPhi amplification were assessed by running 1 μ l of a 1:10 dilution of the GenomiPhi product on a 0.8% agarose gel. Following successful amplification, GenomiPhi products were diluted 1:50 in TE⁻¹ (10mM Tris-HCl pH 7.5, 1mM EDTA pH 8.0) for use in locus-specific PCR reactions.

The dbSNP rs #'s, primers, and amplification conditions for each of the ten SNPs genotyped in this paper are listed in Table 3.3. All SNPs were amplified using 3 μ l of the 1:50 diluted GenomiPhi product in a 25 μ l reaction. Genotyping was performed using the melting curve single nucleotide polymorphism (McSNP) genotyping method (Akey et al., 2001; Ye et al., 2002). Locus-specific PCR and genotyping methods for the geographically diverse

comparison samples were the same as for the Island Melanesian samples. As whole genome amplification was not necessary to generate sufficient quantities of these DNA samples 21ng of genomic DNA from each individual was used in the locus-specific PCR reactions.

Statistical Methods

To control for male-female differences in pigmentation the raw M index values were z-standardized for sex. Using these z-standardized values mean skin and hair pigmentation were compared between the three large island samples (Bougainville, New Britain, and New Ireland), between the two linguistic phyla [Austronesian (AN) and Papuan (P)], and between the 23 large ($n > 10$) neighborhoods using standard ANOVA. As it has been observed (Norton et al. in press) that phylum-level differences were often cross-cut by differences between islands, comparisons of mean pigmentation between linguistic phyla on each island were also examined.

Mean allele frequencies were calculated for each of the five comparison populations and for the island, phylum, and neighborhood categories from the Island Melanesian sample. Allele frequencies were compared between populations using χ^2 tests of two-way contingency tables in the program PopGene v.1.32 (<http://www.ualberta.ca/~efyeh>; Yeh and Boyle, 1997).

Departures from Hardy Weinberg Equilibrium (HWE) for each locus were tested using the Markov Chain method as implemented by the program GenePop (<http://wbiomed.curtin.edu.au/genepop/index.html>).

Table 3.3: Marker amplification and digestion information

Marker	dbSNP rs#	Primer sequence	PCR (°C) ¹	MgCl ₂ (mM)	notes	Enzyme digestion ²	notes
TYR A192C	rs1042602	L-TTATGTGTCAATGGATGCAC R-GCTTCATGGGCAAATCAAT	94/62.5/72	2.5		DpnII	
TYR C308G	rs1042608	L-CCTGGAAACCATGACAAATCt R-CAGGCAAATTTCTACATCAGC	94/54.5/72	2		XbaI	
TYR A402G	rs1800422	L-CAGTATTTTTGAGGCAGTGGCTtC R-TGGCTTCTGGATAAACTTCTTGA	94/56.4/72	2.5		BstBI	65°C digestion
TYRP1 A209T	rs1800374	L-TGGGGTAGGACAGGAAAGC R-GGTACCTGTGCCATGTGAGA	94/51.5/72	1.5		HphI	
OCA2 A355G	rs1800404	L-GCGGTAATTTCTGTGCTTC R-CAGTGCTGCAAGGAACC	94/55/72	1.5		Hae III	
ASIP A8818G	rs6058017	L-CTGCCAGTGCCGCTTCTT R-AAGCCAGGTCTCCCTGAAGT	94/53/72	2	10% DMSO, 1% BSA, 45s	BsrB I	
MATP A272G	rs26722	L-CAGGACCCTCCATTGTCATC R-CCCTGCATTGCCAGCTCT	94/60.4/72	1.5		BsaAI	
MATP C374G	rs16891982	L-CGAAAGAGGAGTCGAGGTTG R-GTGAGGAAAACACGGAGTTGGT	94/56.4/72	2		ApaLI	
MC1R G92A	rs2228479	L-ATCTGCTGCCTGGCCTTG R-TCCAGCAGGAGGATGACG	94/63.4/72	1.5		NlaIII	
MC1R G314A	rs2228478	L-TCATCTACGCCTTCCACAGC R-TGCCCAGCACACTTAAAGC	94/60/72	2.5		Nsp I	

¹ PCR conditions: after initial denaturation at 94°C for 5 minutes, samples were amplified for 30 cycles, followed by a final extension at 72°C for five minutes. All denaturation/annealing/extension times were 30s unless otherwise specified. Amplifications were performed in a 25µl reaction volume with 200µM dNTPs, 10mM Tris-HCl (pH 8.9), 50mM KCl, 1 U Taq polymerase, and 50µM primers

² 10µl of digestion cocktail containing 5 U enzyme were added to 25µl of PCR product and digested at 37°C for 12-24 hours unless otherwise noted

We calculated hierarchical F statistics for the four SNPs that are polymorphic in the Island Melanesian sample (*ASIP* A8818G, *OCA2* A355G, *MC1R* G92A, and *MC1R* G314A) in two different ways. The first considered variation in neighborhoods nested within islands (F_{NI}), islands within the total region (F_{IT}), and neighborhoods within the total region (F_{NT}). If F_{NI} values were consistently greater than those for F_{IT} , then it would demonstrate that within-island heterogeneity in pigmentation alleles was greater than between-island heterogeneity. The second method tried to take into account the within-island heterogeneity in between and within phyla that had been observed for pigmentation phenotype on some islands in Norton et al. Variation between neighborhoods within phyla on a single island (F_{NP}), variation between phyla on a single island (F_{PI}), and neighborhoods within the island as a whole (F_{NI}) were compared for each of the three large islands (New Britain, New Ireland, and Bougainville). If F_{PI} was greater than F_{NI} at a pigmentation locus it would signify that the most important within-island population structuring occurred at phylum level.

By comparing mean skin and hair pigmentation between the three genotype classes at a given locus using standard ANOVA we tested for single-locus effects on mean skin and hair pigmentation. Interaction effects between the four polymorphic loci were also tested for in a similar manner.

Results

Phenotypic Variation

In order to ensure that this sub-sample captured a similar range of variation as the original larger sample set we re-ran analyses of phenotypic variation within and between islands

and phyla (see Table 3.4). Consistent with our observations using the larger sample, heterogeneity in skin and hair pigmentation phenotype exists among islands (skin: $F = 369.95$, $df = 2$, $p < 0.0001$; hair: $F = 5.81$, $df = 2$, $p < 0.01$), as well as among neighborhoods across the region (skin: $F = 33.12$, $df = 22$, $p < 0.0001$; hair: $F = 4.74$, $df = 22$, $p < 0.0001$). Significant differences in hair pigmentation were observed between AN and P speakers pooled across the region ($F = 28.58$, $df = 1$, $p < 0.0001$).

Table 3.4: Results of ANOVAs comparing skin and hair M index between neighborhoods and linguistic phyla on each island

Source	df	Skin		df	Hair	
		F	p		F	p
Bougainville						
between AN and P speakers	1	1.46	0.2296	1	46.33	< 0.0001
among all neighborhoods	3	6.00	0.0009	3	19.04	< 0.0001
among AN neighborhoods	2	6.93	0.0021	2	2.63	0.0850
among P neighborhoods	-	-	-	-	-	-
New Britain						
between AN and P speakers	1	0.23	0.6312	1	5.00	0.0262
among all neighborhoods	11	3.35	0.0002	11	5.15	< 0.0001
among AN neighborhoods	7	2.95	0.0061	7	3.51	0.0016
among P neighborhoods	3	5.21	0.0022	3	9.78	< 0.0001
New Ireland						
between AN and P speakers	1	5.72	0.0181	1	3.39	0.0678
among all neighborhoods	5	1.66	0.1499	5	1.41	0.2255
among AN neighborhoods	2	0.65	0.5276	2	2.21	0.1198
among P neighborhoods	2	1.12	0.3336	2	0.27	0.7643

Previous work using the full pigmentation data set from Norton et al. demonstrated that heterogeneity was present within islands as well as within linguistic phyla. To determine if our sub-sample also captured this heterogeneity mean skin and hair M index values were compared among and between AN and P speaking neighborhoods on each island separately, and between all neighborhoods on the same island. These results are also shown in Table 3.4. As in the larger sample, heterogeneity in pigmentation phenotype exists within islands, as well as within linguistic phyla on the larger islands of Bougainville and New Britain. We observe significant

variation in skin pigmentation on Bougainville among all neighborhoods ($F = 6.00$, $df = 3$, $p < 0.001$) and among the AN-speaking neighborhoods ($F = 6.93$, $df = 2$, $p < 0.01$). Hair pigmentation on Bougainville differs significantly between AN and P speakers ($F = 46.33$, $df = 1$, $p < 0.0001$) as well as among all neighborhoods ($F = 19.04$, $df = 3$, $p < 0.0001$). In New Britain we observe significant variation in skin pigmentation across all neighborhoods ($F = 3.35$, $df = 11$, $p < 0.001$), and among AN ($F = 2.95$, $df = 7$, $p < 0.01$) and P ($F = 5.21$, $df = 3$, $p < 0.01$) speaking neighborhoods. Hair pigmentation on this island was different among all neighborhoods ($F = 5.15$, $df = 11$, $p < 0.0001$), between AN and P speakers ($F = 5.00$, $df = 1$, $p < 0.05$), and among AN ($F = 3.51$, $df = 7$, $p < 0.01$) and P ($F = 9.78$, $df = 3$, $p < 0.0001$). The narrower island of New Ireland is more homogeneous with regards to both skin and hair pigmentation. On this island we only observe a significant difference in skin pigmentation between AN and P speakers ($F = 5.72$, $df = 1$, $p < 0.05$).

Genotypic Variation

Allele frequencies for each of the pigmentation SNPs in each island, neighborhood, and phylum are shown in Table 3.5. Allele frequencies for the five global comparison populations are included for contrast. Six of the SNPs (*TYR* A192C, *TYR* C308G, *TYR* A402G, *TYRPI* A209T, *MATP* A272G, and *MATP* C374G) exhibit quite low levels of polymorphism (allele frequency < 0.01) in the Island Melanesian sample as a whole. Four of these (*TYR* C308G, *TYR* A402G, *TYRPI* A209T, and *MATP* A272G) also show similarly low frequencies in the global population samples. The four SNPs showing appreciable polymorphism (heterozygosity > 0.05) in the Island Melanesian sample are *ASIP* A8818G, *OCA2* A355G, *MC1R* G92A, and *MC1R* G314A.

In Figure 3.1 the allele frequencies of the four polymorphic SNPs in this sample are plotted separately by island, pooled across all Island Melanesians, and for each comparison population. When pooled across all islands Island Melanesians are significantly different ($p < 0.05$) from West Africans and Native Americans at all four loci, from South Asians at *OCA2* A355G, *MC1R* G92A and *MC1R* G314, from East Asians at *ASIP* A8818G, *OCA2* A355G and *MC1R* G92A, and from Europeans at both *MC1R* loci. Within Island Melanesia, Bougainville is significantly different from both New Britain and New Ireland at *ASIP* A8818G and *OCA2* A355G ($p < 0.05$). Although the allele frequencies at these loci in Bougainvilleans are the most similar to those observed in the darkly pigmented West Africans (West African *OCA2* = 0.04, Bougainville *OCA2* = 0.25; West African *ASIP* = 0.15, Bougainville *ASIP* = 0.48), they are still significantly different ($p < 0.05$). The three islands are significantly different from each other at *OCA2* A355G ($p < 0.05$), and New Ireland and New Britain show significant differences in allele frequencies at both *MC1R* loci ($p < 0.05$). The frequencies of the *OCA2* 355*A allele among New Britain islanders (0.64) and among Europeans (0.44) are not significantly different, an interesting fact given that skin pigmentation in this sub-sample was lightest on the island of New Britain.

Table 3.5: Mean allele frequencies according to island, neighborhood, linguistic phylum and comparison samples.

Island	Neighborhood	n	TYR 192*A	TYR 308*C	TYR 402*A	ASIP 8818*A	TYRP1 209*A	OCA2 355*A	MATP 272*A	MATP 374*C	MC1R 92*G	MC1R 314*G
New Britain		288	0.00	0.00	0.01	0.86	0.00	0.64	0.00	0.00	0.86	0.31
	Arimegi Island (Kove)	45	0.00	0.00	0.01	0.85	0.00	0.48	0.00	0.00	0.63	0.43
	Kariai (Anem)	24	0.00	0.00	0.00	0.69	0.00	0.59	0.00	0.00	0.93	0.28
	Pureling (Anem)	27	0.00	0.00	0.00	0.68	0.00	0.45	0.00	0.00	0.92	0.17
	Kisiluvi (Mamousi)	24	0.00	0.00	0.02	0.87	0.00	0.78	0.00	0.00	0.97	0.22
	Lingite (Mamousi)	12	0.00	0.00	0.00	0.83	0.00	0.96	0.00	0.00	1.00	0.09
	Welu (Mamousi)	10	0.00	0.00	0.00	1.00	0.00	0.95	0.00	0.00	1.00	0.11
	Loso (Nakanai)	11	0.00	0.00	0.00	0.95	0.00	0.55	0.00	0.00	0.95	0.30
	Uasilau (Ata)	39	0.01	0.00	0.01	0.97	0.00	0.72	0.00	0.00	0.91	0.32
	Lugei (Ata)	12	0.00	0.00	0.00	1.00	0.00	0.83	0.00	0.00	0.92	0.14
	Bileki (Nakanai)	19	0.00	0.00	0.00	0.97	0.00	0.56	0.00	0.00	0.83	0.41
	Ubili (Melamela)	36	0.00	0.00	0.01	0.84	0.00	0.70	0.00	0.00	0.82	0.36
	Kuanua (Tolai)	15	0.04	0.00	0.00	0.83	0.00	0.40	0.00	0.00	0.82	0.37
New Ireland		158	0.00	0.00	0.01	0.83	0.00	0.45	0.00	0.01	0.79	0.39
	Tigak	25	0.00	0.00	0.02	0.89	0.00	0.70	0.00	0.02	0.63	0.50
	Nailik	21	0.00	0.00	0.00	0.82	0.00	0.53	0.00	0.05	0.75	0.44
	Kabil (Kuot)	33	0.00	0.00	0.05	0.77	0.00	0.27	0.00	0.00	0.83	0.33
	Lamalaua (Kuot)	10	0.00	0.00	0.00	0.90	0.00	0.43	0.00	0.00	1.00	0.17
	Notsi	17	0.00	0.00	0.00	0.81	0.00	0.41	0.00	0.00	0.89	0.22
	Madak	26	0.00	0.00	0.00	0.88	0.00	0.50	0.00	0.00	0.91	0.43
Bougainville		135	0.00	0.00	0.00	0.48	0.01	0.25	0.00	0.00	0.82	0.34
	Saposa Island (Saposa)	36	0.00	0.00	0.00	0.50	0.03	0.22	0.00	0.00	0.79	0.34
	Inivus (Teop)	10	0.00	0.00	0.00	0.43	0.00	0.63	0.00	0.00	0.79	0.43
	Sunahoara (Teop)	10	0.00	0.00	0.06	0.69	0.00	0.22	0.00	0.00	0.94	0.35
	Kukuavo (Aita)	32	0.00	0.00	0.00	0.32	0.00	0.14	0.00	0.00	0.90	0.24
Austronesian		386	0.00	0.00	0.01	0.79	0.00	0.52	0.00	0.00	0.80	0.37
Papuan		219	0.00	0.00	0.01	0.76	0.00	0.46	0.00	0.00	0.90	0.28
East Asian	-	61	0.00	0.00	0.00	0.72	0.00	0.37	0.00	0.02	0.72	0.09
South Asian	-	63	0.04	0.00	0.00	0.77	0.00	0.29	0.00	0.07	0.97	0.19
Native												
American	-	65	0.01	0.00	0.01	0.98	0.00	0.38	0.00	0.06	0.95	0.11
European	-	44	0.52	0.02	0.03	0.86	0.00	0.44	0.00	0.86	0.85	0.10
West African	-	66	0.01	0.00	0.01	0.15	0.01	0.04	0.01	0.05	1.00	0.47

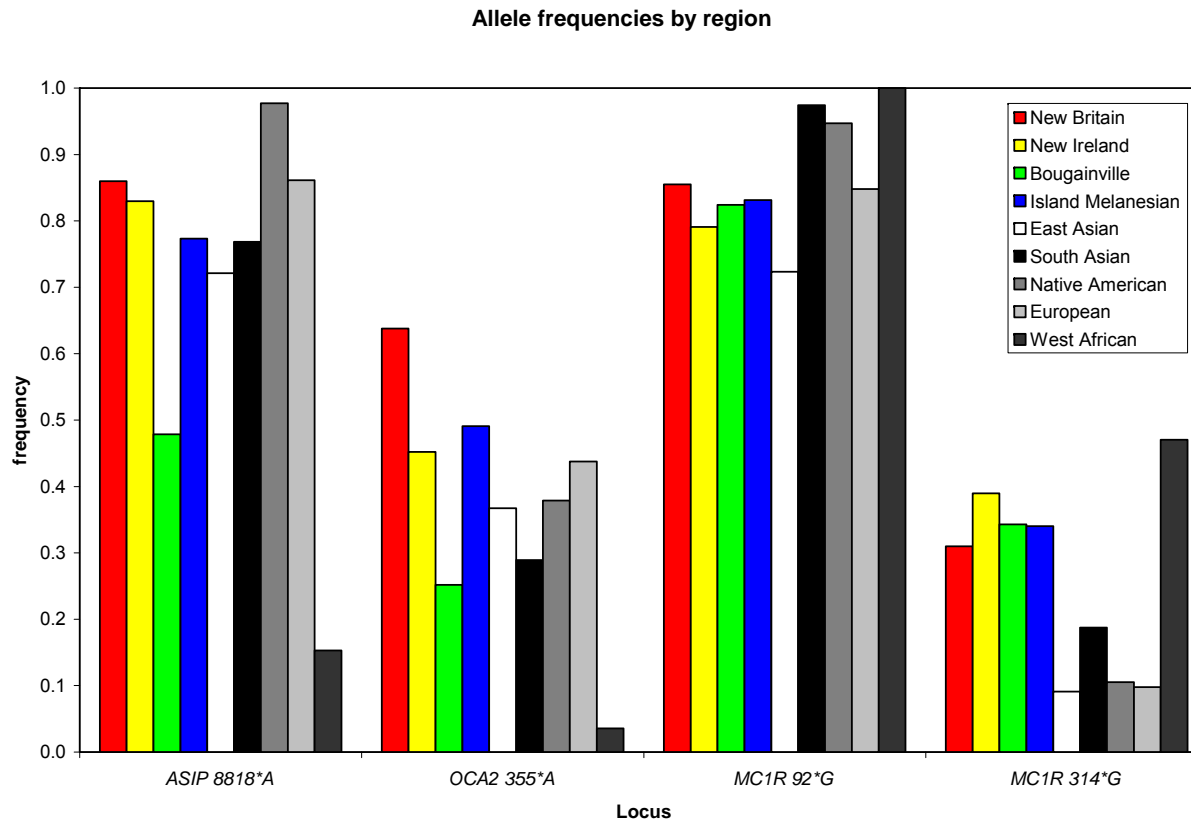


Figure 3.1: Allele frequencies for the four SNPs polymorphic in Island Melanesia according to island, pooled across the region, and in each of the five comparison populations

Hardy Weinberg Equilibrium

Table 3.6 shows the results of Hardy Weinberg tests for a deficiency of heterozygotes at the four polymorphic loci in the total sample, within each island, and within each phylum. In the combined sample, significant departures from HWE were observed at all of the four loci that are polymorphic in our Island Melanesian sample. However, this is not surprising since random mating between islands seems unlikely. Although there is no *a priori* reason to expect homogeneity within islands, it is reasonable to expect to see more homogeneity within islands than between them. Tests of HWE within islands reveal that Bougainville is the only island for which all four SNPs are in HWE. Within New Britain, significant departures from HWE are observed for *OCA2* A355G, *MC1R* G92A, and *MC1R* G314A ($p < 0.05$). On New Ireland, all four polymorphic loci showed significant departures from HWE ($p < 0.05$). This is unexpected given that it is much less geographically complex than Bougainville or New Britain, both of which showed significant intra-island complexity in pigmentation phenotype. Assumptions of random mating within each linguistic phylum on each island are also violated for some loci. On New Britain *OCA2* A355G and *MC1R* G92A shows significant departures from HWE among AN speakers ($p < 0.05$), while at *MC1R* G314A there is a significant deficiency of heterozygotes among Papuan speakers ($p < 0.05$). On New Ireland, *MC1R* G92A is not in HWE ($p < 0.05$) among AN speakers, while *ASIP* A8818G and *OCA2* A355G are not among P speakers ($p < 0.05$).

Table 3.6: Significance tests for departures from Hardy Weinberg Equilibrium (deficiency of heterozygotes)

	ASIP A8818G	OCA2 A355G	MC1R G92A	MC1R G314A
total sample	0.000	0.000	0.000	0.000
AN	0.021	0.000	0.000	0.017
P	0.000	0.000	0.059	0.002
Bougainville	0.367	0.095	0.216	0.301
AN within Bougainville	0.808	0.229	0.232	0.548
P within Bougainville	0.154	0.227	1.000	0.481
New Britain	0.055	0.035	0.002	0.037
AN within New Britain	0.276	0.009	0.008	0.276
P within New Britain	0.071	0.580	0.407	0.022
New Ireland	0.017	0.024	0.001	0.016
AN within New Ireland	0.532	0.602	0.019	0.119
P within New Ireland	0.001	0.018	0.060	0.092

Hierarchical Locus-Specific F Statistics

Hierarchical locus-specific F_{ST} was first calculated for neighborhoods within islands, islands within the Island Melanesian region as a whole, and neighborhoods across Island Melanesia. These F_{ST} values for the four polymorphic SNPs are presented in Table 3.7. Variation between islands at the *ASIP* A8818G locus was greater ($F_{IT} = 0.13$) than the variation found within islands ($F_{NI} = .04$). This difference could be related in part to the sharp differences in pigmentation between Bougainville and the other islands. It is interesting that we do not see a similar result for *OCA2* A355G, a locus at which all three islands were significantly different from each other. It should be noted that the difference in within-island vs. between island variation at this locus was small ($F_{NI} = 0.10$, $F_{IT} = .08$). At both *MC1R* loci variation within islands was greater than variation observed between islands, although in both cases the locus-specific F-statistics were low (≤ 0.06), implying little divergence between or within islands at these SNPs. Comparisons of variation among neighborhoods within phyla on each island reveal

that on the large island of New Britain only *OCA2* A355G shows higher variation between phyla ($F_{PI(OCA2)} = 0.09$) than within ($F_{NP(OCA2)} = 0.04$). Conversely, variation between phyla on Bougainville was greater than variation within phyla at all loci except for *OCA2* A355G ($F_{NP(OCA2)} = 0.10$; $F_{PI(OCA2)} = 0.04$). New Ireland showed relatively low levels of variation across all loci, both within and between phyla.

Table 3.7: Hierarchical F_{ST} values in the region at large, and within each island

	F_{NI}	F_{IT}	F_{NT}
Total Region			
<i>ASIP</i> A8818G	0.04	0.13	0.17
<i>OCA2</i> A355G	0.10	0.08	0.17
<i>MC1R</i> G92A	0.06	0.01	0.07
<i>MC1R</i> G314A	0.02	0.01	0.03
New Britain	F_{NP}	F_{PI}	F_{NI}
<i>ASIP</i> A8818G	0.15	0.03	0.18
<i>OCA2</i> A355G	0.04	0.09	0.12
<i>MC1R</i> G92A	0.08	0.00	0.08
<i>MC1R</i> G314A	0.03	0.00	0.04
New Ireland	F_{NP}	F_{PI}	F_{NI}
<i>ASIP</i> A8818G	0.00	0.00	0.00
<i>OCA2</i> A355G	0.02	0.02	0.04
<i>MC1R</i> G92A	0.04	0.05	0.08
<i>MC1R</i> G314A	0.03	0.01	0.03
Bougainville	F_{NP}	F_{PI}	F_{NI}
<i>ASIP</i> A8818G	0.00	0.04	0.03
<i>OCA2</i> A355G	0.10	0.04	0.13
<i>MC1R</i> G92A	0.00	0.01	0.00
<i>MC1R</i> G314A	0.00	0.01	0.00

Genotype-phenotype Associations

We tested for single-locus effects on pigmentation phenotype using standard ANOVA.

The results are shown in Table 3.8, both for analyses using data pooled across the total region

as well as for separate island analyses. When populations are pooled across the region there is a significant association between *ASIP* A8818G genotype and skin pigmentation ($F = 37.16$, $df = 2$, $p < 0.0001$), *OCA2* A355G genotype and skin pigmentation ($F = 40.97$, $df = 2$, $p < 0.0001$) and *MC1R* G314A genotype and hair pigmentation ($F = 3.35$, $df = 2$, $p < 0.0358$). However the analyses of the phenotypic data revealed that the sample is heterogeneous for pigmentation phenotype, particularly between islands. It was also observed that allele frequencies at two of these loci (*OCA2* A355G and *ASIP* A8818G) differ significantly between islands. This raises an interesting question: are the associations between genotype and phenotype that are observed here measuring actual functional effects, or might they instead be the result of population stratification? To determine if similar effects could also be observed within subpopulations, we tested for genotype-phenotype associations one island at a time. In these analyses, *OCA2* A355G has a significant effect on skin pigmentation only in New Britain ($F = 3.03$, $df = 2$, $p < 0.0498$), and *MC1R* G314A has a significant effect on hair pigmentation only in New Ireland ($F = 5.35$, $df = 2$, $p < 0.0059$).

We also tested for interaction effects between candidate SNPs. Significant results of interactions are reported in Table 3.9. Significant interaction effects between *ASIP* A8818G and *OCA2* A355G on skin pigmentation are observed in the sample at large, among Papuan speakers across all islands, and among Papuan speakers on New Ireland. *OCA2* A355G and *MC1R* G314A have a significant interaction effect on skin pigmentation in the sample at large, and on skin and hair pigmentation within Bougainville. Together the two *MC1R* loci have a significant interaction effect on Papuan hair pigmentation.

Table 3.8: Genotype-phenotype associations within the sample-at-large, and on each island

Source	df	skin		df	Hair	
		F	p		F	p
Total region						
ASIP A8818G	2	37.16	< 0.0001	2	2.99	0.0510
OCA2 A355G	2	40.97	< 0.0001	2	2.03	0.1320
MC1R G92A	2	0.56	0.5737	2	2.38	0.0940
MC1R G314A	2	0.62	0.5383	2	3.35	0.0358
Bougainville						
ASIP A8818G	2	0.57	0.5665	2	1.51	0.2266
OCA2 A355G	2	2.9	0.0590	2	0.92	0.4011
MC1R G92A	2	0.14	0.8653	2	1.13	0.3286
MC1R G314A	2	0.05	0.9559	2	0.42	0.6569
New Britain						
ASIP A8818G	2	1.76	0.1746	2	0.05	0.9468
OCA2 A355G	2	3.03	0.0498	2	0.04	0.9656
MC1R G92A	2	1.83	0.1627	2	0.49	0.6126
MC1R G314A	2	0.59	0.5560	2	0.83	0.4378
New Ireland						
ASIP A8818G	2	1.22	0.2974	2	0.94	0.3921
OCA2 A355G	2	1.76	0.1756	2	0.84	0.4319
MC1R G92A	2	0.42	0.6556	2	1.59	0.2085
MC1R G314A	2	0.09	0.9132	2	5.35	0.0059

Table 3.9: Significant secondary interaction effects on skin and hair pigmentation

level	interaction	F	Skin		F	Hair	
			df	p		df	p
at large	ASIP A8818G*OCA2 A355G	2.46	4	0.0441	-	-	-
P	ASIP A8818G*OCA2 A355G	3.18	4	0.0148	-	-	-
P NI	ASIP A8818G*OCA2 A355G	-	-	-	3.56	3	0.0208
at large	OCA2 A355G*MC1R G92A	2.64	4	0.0333	-	-	-
at large	OCA2 A355G*MC1R G314A	2.55	4	0.0382	-	-	-
B	OCA2 A355G*MC1R G314A	3.27	4	0.0148	3.22	4	0.0171
P	MC1R G92A*MC1R G314A	-	-	-	3.30	2	0.0395

Discussion

Phenotypic Variation

As observed in the larger Island Melanesian sample, the subset of Island Melanesians used in these genetic analyses shows significant heterogeneity in skin and hair pigmentation. Primary distinctions are found among islands and among neighborhoods within those islands. This is important for a number of reasons. First, as observed in Chapter Two and Norton et al. these findings demonstrate that we cannot assume a basic general pigmentation “type” for all Island Melanesians. Understanding that such heterogeneity exists can be important when attempting to detect associations between genotype and phenotype because failure to control for those associations may result in false-positives. Such false-positives are more likely to occur when the marker being tested also shows variation between the populations of interest, as in the case here. As the variation in pigmentation cannot be attributed to solely to the effects of differential UVR across islands, this suggests that either genetic drift or sexual selection (or both) may have played in important role in shaping genetic variation in these populations.

Genotypic Variation

The ten SNPs genotyped in this study were selected based on either reports in the literature suggesting associations with skin pigmentation variation (often between populations of European and West African descent), allele frequency differences between populations known to differ in pigmentation phenotype, or because of their location in pigmentation candidate genes.

A description of the individual SNPs, and the individual pigmentation genes in which they are found is provided in Chapter One.

The lack of variation at six of the SNPs in the Island Melanesian sample warrants discussion. Two of the SNPs that showed little to no variation were *TYR* A192C and *MATP* C374G. Results presented in Chapter Four show that the *TYR* 192*A and *MATP* 374*C alleles are common almost exclusively in European populations and are associated with lighter skin pigmentation. The lack of variation in Island Melanesia at these loci may be because the *TYR* 192*A and *MATP* *374 C alleles are associated with lighter pigmentation and/or because they appear to be relatively recent mutations. The remaining four SNPs that show low polymorphism in Island Melanesia (*TYR* C308G, *TYR* A402G, *TYRP1* A209T, and *MATP* A272G) also exhibit low variation in the comparison populations, although reports from dbSNP indicated that they were polymorphic. This raises the important issue of the effects that ascertainment bias in SNP discovery can have on studies such as this. Many SNPs available in public databases have been typed in a small number of chromosomes and in a limited sample of populations, usually of European origin (Akey et al., 2003). It may be that while these SNPs showed some variation in European-derived populations they are not as variable outside of Europe. This does not mean that the genes considered in this study have no effect on pigmentation variation in Island Melanesia; rather, it suggests that SNPs in these genes that have been primarily discovered in European populations may not be useful in explaining pigmentation variation in Island Melanesia (or other regions). One method to overcome this ascertainment bias would be to sequence a small Island Melanesian sample for this same set of genes (or selected coding regions within each gene) to identify any polymorphisms that might be common.

While ascertainment bias may contribute in part to the low polymorphism levels observed in the Island Melanesian sample, low polymorphism in the region is not necessarily surprising when we consider that high UVR across the region may have constrained genetic variation at some loci via natural selection. This constraint would be especially strong on loci that have a major effect on lowering skin pigmentation levels; the low frequency of *TYR* 192*A and *MATP* 374*C may be examples of this. It is possible that these two SNPs are in fact affecting overall pigmentation in the region (causing darker skin pigmentation), but that they are not responsible for within-population variation. East Asians also have the non-European alleles for both of these loci at high frequency, suggesting that perhaps East Asians are lightly pigmented due to alleles at other pigmentation loci. In this high UVR environment natural selection may favor polymorphisms that prevent pigmentation from dropping below a certain threshold level (as discussed in Chapter Two). Mutations that are associated with fluctuation above this level would not be considered deleterious, and so might increase in frequency in these populations due to either drift or sexual selection. As some of the pigmentation loci that were typed in this study are associated with major differences in pigmentation between populations they may not be the best candidates to explain the comparatively minor differences in pigmentation that we observed between Island Melanesian populations.

Two of the four SNPs identified as polymorphic in Island Melanesia are associated with normal variation between European and African Americans--*ASIP* A8818G and *OCA2* A355G (Bonilla et al. 2005, Shriver et al. 2003). All three islands were significantly different from each other at *OCA2* A355G and that Bougainville was significantly different from New Britain and New Ireland at *ASIP* A8818G. The other two polymorphic loci are both in the *MC1R* gene, but it has not been demonstrated that either has an effect on normal variation between Europeans and

Africans. Although the *MC1R* G92A variant is considered a “weak” red hair color allele (Duffy et al. 2004), it is also polymorphic among East Asian populations (Rana et al., 1999; Yao et al., 2000; Peng et al., 2001). As many Island Melanesians are descended from migrants arriving as part of the Austronesian expansion from insular southeast Asia we expected to observe polymorphism at this locus. All three islands were significantly different from each other at *OCA2* A355G and that Bougainville was significantly different from New Britain and New Ireland at *ASIP* A8818G. New Britain and New Ireland were significantly different from each other at both *MC1R* loci. The association of this variation with phenotype will be discussed below.

In addition to comparing allele frequencies among the islands in our study, it is also instructive to compare the allele frequencies to other populations from around the world. When the Island Melanesians are pooled together they are significantly different from West Africans and Native Americans at all four of the polymorphic SNPs. The differences between both groups may be due to the relatively long period of time since each shared a common ancestor with Island Melanesians. Given the relatively dark skin observed across this Island Melanesian sample we might expect pigmentation allele frequencies to be most similar to those of West Africans, whose skin pigment is also relatively dark. While Island Melanesians as a group are more similar in allele frequencies to West Africans than to the other populations at the *ASIP* A8818G and *OCA2* A355G loci, these differences in allele frequencies are still statistically significant.

Hardy Weinberg Equilibrium

Although we observed significant departures from HWE at all four of the SNPs that were polymorphic in the Island Melanesian sample at large, by pooling populations across islands we may have artificially created population substructure. If so, then the significant deficiency of heterozygotes could be the result of the Wahlund effect. One possible way to test for this would be to examine HWE in each subpopulation, or island, separately. We observe that the islands of New Britain and New Ireland show departures from HWE while Bougainville did not. These HWE deviations at the island level may be because of within island population substructure. For example, New Britain is a large island with a relatively mountainous interior. It is possible that the topography of the island, as well as its overall size, may have presented some barrier to random mating among populations on the island. New Ireland, however, is a smaller, much narrower island, and potential geographical barriers to gene flow here are less obvious, making the deviations from HWE here more difficult to explain. The rugged terrain of Bougainville might present some geographical impediments to random mating, but as our sample was from primarily the northern region of the island this may have limited relevance.

Hierarchical Locus-Specific F Statistics

Analyses of pigmentation variation identified significant heterogeneity within and between islands in the Island Melanesian sample. Comparisons of hierarchical F statistics for pigmentation candidate SNPs within and between islands, phyla and neighborhoods suggest a similar complexity for *ASIP* and *OCA2* while both *MC1R* loci show little hierarchical stratification. The motive for calculating these hierarchical F statistics was to be able to

determine what level of variation was most important in shaping heterogeneity in pigmentation candidate genes in the region. For example, if F_{IT} values were consistently higher than F_{NI} values then it would indicate that between-island differences were more important than within island differences. As only four loci were polymorphic in the sample it is difficult to discuss broad trends for pigmentation loci in terms of these F statistics. Three of the four loci (*OCA2* A355G, *MC1R* G92A, and *MC1R* G314A) showed higher levels of variation among neighborhoods within island than between islands (although F_{NI} and F_{IT} values for *OCA2* A355G were similar). However, the values of these F statistics are generally low, indicating little genetic variation among populations.

Similar calculations were performed to compare variation within and between phyla on each island. Again, due to the small number of loci it is difficult to explore general trends in these values. As such, it may be more interesting to consider overall patterns within each island. New Ireland shows low (≤ 0.05) F_{NP} and F_{PI} values across all loci, while Bougainville shows similarly low values at all loci except for *OCA2* A355G, where $F_{NP} = 0.10$, suggesting moderate genetic differentiation among neighborhoods within phyla. On New Britain F_{NP} and F_{PI} values are on average higher than is observed on other islands. Variation among neighborhoods within phyla is greater than between phyla at *ASIP* A8818G, but unlike Bougainville it is lower at *OCA2* A355G. New Britain also has the highest F_{NP} values for both *MC1R* loci of any of the islands ($F_{NP\ G92A} = 0.08$, $F_{NP\ G314A} = 0.03$).

Inferences about the significance of these hierarchical locus-specific F-statistics are limited by our lack of knowledge of the average genetic variation between islands and phyla at neutral loci. For example, is the presumably high F_{IT} value at *ASIP* A8818G of 0.13 truly elevated (indicating strong inter-island divergence at this locus), or is it simply consistent with

average genetic differences between the three islands? Questions like this will be better resolved when we know the average F_{IT} (or F_{NI} , F_{PI}) value across a large number of neutral autosomal loci.

Genotype-phenotype Associations

Tests for single-locus effects on skin or hair pigmentation across the pooled sample reveal significant associations between skin pigmentation and genotype at *OCA2* A355G and *ASIP* A8818G, and between hair pigmentation and *MC1R* G314A genotype. However, because we observe significant differences between islands in pigmentation phenotype and in allele frequencies at these loci it is difficult to determine if these results represent true associations or false positives resulting from cryptic stratification. In an effort to overcome possible issues of population stratification we tested for genotype-phenotype associations one island at a time. In this case, we find only significant associations between *OCA2* A355G and skin pigmentation on New Britain and *MC1R* G314A on hair pigmentation on New Ireland. Of course, by testing for effects within each island our sample size is reduced, which will decrease the power to detect an association. Significant interactions between *OCA2* A355G and *MC1R* G314A were observed on Bougainville for both skin and hair pigmentation. Interactions between other *OCA2* and *MC1R* loci on normal and albino phenotypes have been previously reported in other populations. Akey et al. (2001) demonstrated a significant interaction effect between an *OCA2* SNP (IVS13-15) and *MC1R* G92A in a Tibetan population (Akey et al., 2001) and Duffy et al. observed interaction effects on skin pigmentation, freckling, and nevus count among individuals of European descent (Duffy et al., 2004). Interaction effects between *OCA2* and *MC1R* loci were also observed on the oculocutaneous albinism type two phenotype (King et al., 2003).

We feel that the strong associations between skin pigmentation and genotype at *ASIP* A8818G and *OCA2* A355G in the pooled sample may be influenced by population substructure. This stratification could be the result of admixture between subpopulations heterogeneous for pigmentation phenotype, which can lead to a false positive genotype association when allele frequencies also vary between subpopulations (Hoggart et al., 2003). By genotyping an individual for marker loci that show large differences between subpopulations (ancestry informative markers, or AIMs) it is possible to control for the confounding effects of substructure (Chakraborty, 1975). While the Island Melanesians included in this sample have been extensively studied for Y chromosome and mtDNA variation, less is currently known about variation at autosomal loci. It is hoped that future typing of large numbers of autosomal loci (as on the Affymetrix 10K WGA Gene Chip) in these populations may allow for differentiation between different Island Melanesian populations. This may make it possible to apply admixture mapping techniques to tease apart effects of population structure and to detect true genotype-phenotype associations.

With this in mind there are two major difficulties in applying admixture mapping techniques to Island Melanesia that must be explored: the lack of clearly defined “ancestral” populations and the length of time since admixture occurred. Many admixture mapping studies deal with African American or Hispanic populations, for which ancestral populations can usually be clearly identified and time since admixture is relatively recent (Shriver et al., 1997; Parra et al., 1998; Molokhia and McKeigue, 2000; Smith et al., 2001; Fernandez et al., 2003; Shriver et al., 2003; Bonilla et al., 2004; Collins-Schramm et al., 2004). Assuming that the two “ancestral” populations contributing to modern-day Island Melanesians were the Austronesian speakers associated with Lapita culture and the progenitors of modern Papuan speaking groups (possibly

some of the earliest arrivals to the region) we are faced with a number of problems regarding the identification of suitable modern populations to use as proxies for ancestral Island Melanesian groups. While some East Asian populations (particularly Taiwanese) may represent an appropriate “Austronesian” ancestral population, there is no group that stands out as a suitable “Papuan” ancestral population. This is most likely due to the fact that Papuan speakers may be the descendants of groups that arrived in Island Melanesia beginning ~ 40 KYA and who spent much of that time relatively (but not completely) isolated from each other. These extended periods of isolation may have caused these early settler populations to diverge from one another due to genetic drift. Some hypotheses of human dispersal suggest an early migration (~ 65 KYA) out of Africa that followed the coasts of southern India and southeast Asia into Oceania (Nei and Roychoudhury, 1993; Cavalli-Sforza et al., 1994; Quintana-Murci et al., 1999; Macaulay et al., 2005). While modern Papuan groups could be descended from these early migrants there is little evidence suggesting where they may have come from or how many waves of migrations like this entered Oceania.

The second potential problem in applying traditional admixture mapping approaches to Island Melanesian populations is that the admixture in these groups is much older than that the types of populations that are commonly used in admixture mapping studies. Admixture between the ancestors of AN and P speaking groups likely did not occur before 3,200 years ago, and there may be even older admixture events that occurred between different P speaking groups across the region. The effectiveness of admixture mapping in cases where the admixture event is this old is unknown.

Conclusions

Island Melanesians are an extremely diverse human population, both linguistically and biologically. As a group they also show a remarkable amount of phenotypic diversity in skin and, to a lesser extent, hair pigmentation. The purpose of this paper was to screen ten pigmentation candidate SNPs in the hopes of identifying genotype-phenotype associations in Island Melanesians. Few such associations are observed, and those that are remain difficult to interpret due to population stratification in the region.

Our findings may have been limited in part by the SNPs that we chose to examine. Few of these pigmentation candidate genes have been screened in Island Melanesians, so the primary allele frequency information in the literature has been based on studies of European, African, Native American, and East Asian populations. Thus, while the SNPs that we selected show variation in other regions of the world, they may be notably less polymorphic in Island Melanesia. Direct sequencing of a sample of Island Melanesians might identify SNPs in these candidate genes that are polymorphic and do have an effect on phenotype in Island Melanesia. Also, we only sampled ten SNPs—there are many more that have been recently identified or genotyped in public databases (e.g. HapMap at www.hapmap.org) that may show variation. It is also possible that SNPs in other pigmentation candidate genes affect pigmentation in Island Melanesians.

While we felt that many of the SNPs that we typed in this study had the potential to be associated with normal variation in pigmentation in Island Melanesia due to previously reported admixture mapping linkage in African-Americans and African-Caribbeans, it is possible that those SNPs are primarily responsible for major effects on pigmentation, and so might not be appropriate to explain the more subtle (although still notable) variation observed among different

Island Melanesian populations. Two loci that are not polymorphic among Island Melanesians but that may still be having a strong effect on pigmentation are *TYR* A192C and *MATP* C374G. The *TYR* 192*A allele and the 374*C variant in *MATP* are only found at very high frequencies in Europeans, suggesting that they may contribute to overall lighter pigmentation in European and European-derived populations. Island Melanesians are at or near fixation for the non-European allele. Two SNPs that have shown admixture mapping signals affecting normal pigmentation variation between Europeans and West Africans (*ASIP* A8818G and *OCA2* A355G) do show variation within this population, and there is some evidence for an association with normal phenotypic variation for *OCA2* on New Britain. The strength of this association is questionable because of the underlying genetic heterogeneity in the region.

While our knowledge of settlement and some migration patterns into and through Island Melanesia has been greatly improved with help from archaeology (White, 1972; Marshall and Allen, 1991; Summerhayes and Allen, 1993; Pavlides and Gosden, 1994; Spriggs, 1995; Allen, 1996; Spriggs, 1997; Leavesley and Allen, 1998; Leavesley et al., 2002; Summerhayes, in preparation) linguistics (Ross, 1994; Dutton, 1995), and molecular anthropology (Giles et al., 1965; Friedlaender and Steinberg, 1970; Serjeantson and Board, 1992; Serjeantson and Gao, 1995; Merriwether et al., 1999; Redd and Stoneking, 1999; Capelli et al., 2001; Kayser et al., 2001; Kayser et al., 2003; Robledo et al., 2004; Friedlaender et al., 2005) we know much less about localized migration and mating patterns within the region and the effects that these have had on population stratification. Presumably the geographic barriers within and between islands would have presented at least some hindrance to random mating even though archaeological evidence indicates trade between regions as early as 20,000 years BP (Marshall and Allen, 1991; Summerhayes and Allen, 1993; Allen, 1996; Leavesley and Allen, 1998; Summerhayes, in

preparation). This partial isolation could have led to the divergence of populations on different islands at both pigmentation genes and other loci. It is likely that variation at pigmentation candidate loci in this region has been broadly constrained by natural selection (the “melanin threshold” model), but genetic drift and sexual selection have also contributed as well to the complexity observed in this region. By developing methods to control for population stratification in the region it may be possible to more accurately identify true genotype-phenotype associations for pigmentation as well as for other phenotypic traits.

Literature cited

- Akey J, Sosnoski D, Parra E, Dios S, Hiester K, Su B, Bonilla C, Jin L, and Shriver M (2001) Melting curve analysis of SNPs (McSNP): a simple gel-free low cost approach to SNP genotyping and DNA framgnet analysis. *Biotechniques* 30:358-367.
- Akey J, Wang H, Xiong M, Wu H, Weida L, Shriver M, and Jin L (2001) Interaction between the melanocortin-1 receptor and *P* genes contributes to inter-individual variation in skin pigmetnation phenotypes in a Tibetan population. *Human Genetics* 108:516-520.
- Akey J, Zhang K, Xiong M, and Jin L (2003) The effect of single nucleotide polymorphism identification strategies on estimates of linkage disequilibrium. *Molecular Biology and Evolution* 20:232-242.
- Allen J (1996) The Pre-Austronesian settlement of Island Melanesia: Implications for Laptia Archaeology. In W Goodenough (ed.): Prehistoric settlement of the Pacific. Philadelphia: Trasactions of the American Philosophical Society 86, pp. 11-27.
- Bonilla C, Boxill L, Donald S, Williams T, Sylvester N, Parra E, Dios S, Norton H, Shriver M, and Kittles R (2005) The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Human Genetics* 116:402-406.
- Bonilla C, Parra E, Pfaff C, Dios S, Marshall J, Hamman R, Ferrell R, Hoggart C, McKeigue P, and Shriver M (2004) Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Annals of Human Genetics* 68 (*Pt* 2):139-153.

- Branda R, and Eaton J (1978) Skin color and nutrient photolysis: an evolutionary hypothesis. *Science* 201:625-626.
- Capelli C, Wilson J, Richards M, Stumpf M, Garrix F, Oppenheimer S, Underhill P, Pascali V, Tasang-Ming K, and Goldstein D (2001) A predominantly indigenous paternal heritage for the Austroensian-Speaking peoples of insular Southeast Asia and Oceania. *American Journal of Human Genetics* 68:432-443.
- Cavalli-Sforza L, Menozzi P, and Piazza A (1994) *The History and Geography of Human Genes*. Princeton: Princeton University Press.
- Chakraborty R (1975) Estimation of race admixture: a new method. *American Journal of Physical Anthropology* 42:507-511.
- Chaplin G (2004) Geographic distribution of environmental factors influencing human skin coloration. *American Journal of Physical Anthropology* 125:292-302.
- Collins-Schramm H, Chima B, Morii T, Wah K, Figueroa Y, Criswell L, Hanson R, Knowler W, Silva G, Belmont J, and Seldin M (2004) Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians. *Human Genetics* 114:263-271.
- Cowles R (1959) Some ecological factors bearing on the origin and evolution of pigment in the human skin. *American Naturalist* 93:283-293.
- Darwin C (1871) *The Descent of Man, and Selection in Relation to Sex*. London: J. Murray.
- Davis M, Kavita U, Li W, Dar M, Nelson J, Champion M, Jones K, Deadman R, Rhee S, and Mamone J (2002) GenomiPhi whole genome amplification - Amplification of genomic

- DNA from a limited source by Phi29 polymerases for use in genetic analysis: The World Genomics Symposium and Exhibition. Atlantic City, NJ.
- Dean F, Nelson J, Giesler T, and Lasken R (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *11*:1095-1099.
- Diamond J (1988) Survival of the Sexiest. *Discover May*:74-81.
- Diamond J (1992) *The Third Chimpanzee*. New York: Harper Collins Publishers.
- Duffy D, Box N, Chen W, Palmer J, Montgomery G, James M, Hayward N, Martin N, and Sturm R (2004) Interactive effects of MC1R and OCA2 on melanoma risk phenotypes. *Human Molecular Genetics 13*:447-461.
- Dutton T (1995) Language contact and change in Melanesia. In P Bellwood, J Fox and D Tryon (eds.): *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University, pp. 17-38.
- Fernandez J, Shriver M, Beasley T, Rafla-Demetrious N, Parra E, Albu J, Nicklas B, Ryan A, McKeigue P, Hoggart C, Weinsier R, and Allison D (2003) Association of African genetic admixture with resting metabolic rate and obesity among women. *Obesity Research 11*:904-911.
- Friedlaender J (1975) *Patterns of Human Variation: The demography, genetics, and phenetics of Bougainville Islanders*. Cambridge: Harvard University Press.
- Friedlaender J, Schurr T, Gentz F, G K, Friedlaender F, Horvat G, Babb P, Cerchio S, Kaestle F, Schanfield M, Deka R, Yanagihara R, and Merriwether D (2005) Expanding Southwest Pacific mitochondrial haplogroups P and Q. *Molecular Biology and Evolution in press*.

- Friedlaender J, and Steinberg G (1970) Anthropological significance of gamma globulin (Gm and Inv) antigens in Bougainville Island, Melanesia. *Nature* 228:59-61.
- Frost P (1988) Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine* 32:38-58.
- Giles E, Ogan E, and Steinberg A (1965) Gamma-globulin factors (Gm and Inv) in New Guinea: anthropological significance. *Science* 150:1158-1160.
- Hoggart C, Parra E, Shriver M, Bonilla C, Kittles R, Clayton D, and McKeigue P (2003) Control of confounding of genetic associations in stratified populations. *American Journal of Human Genetics* 72:1492-1504.
- <http://www.ualberta.ca/%7efyeh>.
- Jablonski N, and Chaplin G (2000) The evolution of human skin coloration. *Human Evolution* 39:57-106.
- Kayser M, Brauer S, Weiss G, Schiefenhovel S, Underhill P, Shen P, Oefner P, Tommaseo-Ponzetta M, and Stoneking M (2003) Reduced Y-chromosome, but not mitochondrial DNA, diversity in human populations from West New Guinea. *American Journal of Human Genetics* 72:281-302.
- Kayser M, Brauer S, Weiss G, Shiefenhovel W, Underhill P, and Stoneking M (2001) Independent histories of human Y chromosomes from Melanesia and Australia. *American Journal of Human Genetics* 68:173-190.
- King R, Willaert R, Schmidt R, Pietsch J, Savage S, Brott M, Fryer J, Summers C, and Oetting W (2003) MC1R mutations modify the classic phenotype of oculocutaneous albinism type 2 (OCA2). *The American Journal of Human Genetics* 73:638-645.

- Leavesley M, and Allen J (1998) Dates, disturbance and artefact distributions: Another analysis of Buang Merabek, a Pleistocene site on New Ireland, Papua New Guinea. *Archaeology in Oceania* 33:63-82.
- Leavesley M, Bird M, Fifield L, Hausadaen P, Santos G, and di Tada M (2002) Buang Merabek: Early evidence for human occupation in the Bismarck Archipelago, Papua New Guinea. *Australian Archaeology* 54:55-57.
- Lizardi P, Huang X, Zhu Z, Bray-Ward P, Thomas D, and Ward D (1998) Mutation detection and single-molecule counting using isothermal rolling-circle amplification. *Nature Genetics* 19:225-232.
- Loomis W (1967) Skin pigment regulation of vitamin-D biosynthesis in man. *Science* 157:501-506.
- Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F, Taha A, Kassim Shaari N, Maripa Raja J, Ismail P, Zainuddin Z, Goodwin W, Bulbeck D, Bandelth H, Oppenheimer S, Torroni A, and Richards M (2005) Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308.
- Mackintosh J (2001) The antimicrobial properties of melanocytes, melanosomes, and melanin and the evolution of black skin. *Journal of Theoretical Biology* 211:101-113.
- Marshall B, and Allen J (1991) Excavations at Panakiwuk Cave, New Ireland. In J Allen and C Gosden (eds.): Report of the Lapita Homeland Project. Canberra: Research School of Pacific Studies, pp. 59-91.

- Merriwether D, Friedlaender J, Mediavilla J, Mgone C, Gentz F, and Ferrell R (1999) Mitochondrial DNA variation is an indicator of Austronesian influence in Island Melanesia. *American Journal of Physical Anthropology* 110:243-270.
- Molokhia M, and McKeigue P (2000) Risk for rheumatic disease in relation to ethnicity and admixture. *Arthritis Research* 2:115-125.
- Nei M, and Roychoudhury A (1993) Evolutionary relationships of human populations on a global scale. *Molecular Biology and Evolution* 10:927-943.
- Parra E, Kittles R, Argyropoulos G, Pfaff C, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey W, Jin L, McKeigue P, Kamboh M, Ferrell R, Pollitzer W, and Shriver M (2001) Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *American Journal of Physical Anthropology* 114:18-29.
- Parra E, Marcini A, Akey J, Martinson J, Batzer M, Cooper R, Forrester T, Allison D, Deka R, Ferrell R, and Shriver M (1998) Estimating African American admixture proportions by use of population-specific alleles. *American Journal of Human Genetics* 63:1839-1851.
- Pavlidis C, and Gosden C (1994) 35,000 year-old sites in the rain forests of West New Britain, Papua New Guinea. *Antiquity* 69:605-610.
- Peng S, Lu X, Luo H, Xinag-Yu J, and Zhang Y (2001) Melanocortin-1 receptor gene variants in four Chinese ethnic populations. *Cell Research* 11:81-84.
- Post P, Daniels F, and Binford R (1975) Cold injury and the evolution of "white" skin. *Human Biology* 47:65-80.

- Quintana-Murci L, Semino O, Bandelt H, Passarino G, McElreavey K, and Santachiara-Benerecetti A (1999) Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nature Genetics* 23:437-441.
- Rana B, Hewett-Emmett D, Jin L, Chang B, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde L, Ramsay M, Jenkins T, and Li W (1999) High polymorphism at the human melanocortin 1 receptor locus. *Genetics* 151:1547-1557.
- Redd A, and Stoneking M (1999) Peopling of Sahul: mtDNA variation in Aboriginal Australian and Papua New Guinean populations. *American Journal of Human Genetics* 65:808-828.
- Relethford J (1997) Hemispheric difference in human skin color. *American Journal of Physical Anthropology* 104:449-457.
- Robledo R, Scheinfeldt L, Merriwether D, Thompson F, and Friedlaender J (2004) Pattern of Genetic Diversity in Island Melanesia. *Human Biology* 75.
- Ross M (1994) Areal phonological features in north central New Ireland. In T Dutton and D Tryon (eds.): *Language Contact and Change in the Austronesian World*. New York: Mouton de Gruyter, pp. 551-572.
- Serjeantson S, and Board B (1992) Population genetics in Papua New Guinea. In R Attenborough and M Alpers (eds.): *Human biology in Papua New Guinea: The small cosmos*. Oxford: Oxford Science Publications, pp. 198-233.
- Serjeantson S, and Gao X (1995) *Homo sapiens* is an evolving species: Origins of the Austronesians. In P Bellwood, J Fox and D Tryon (eds.): *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University, pp. 17-38.
- Shriver M, Parra E, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R

- (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Human Genetics* 112:387-399.
- Shriver M, Smith M, Jin L, Marcini A, Akey J, Deka R, and RE F (1997) Ethnic-affiliation estimation by use of population-specific DNA markers. *American Journal of Human Genetics* 60:957-964.
- Smith M, Lautenberger J, Shin H, Chretien J, Shrestha S, Gilbert D, and O'Brien S (2001) Markers for mapping by admixture linkage disequilibrium in African American and Hispanic populations. *American Journal of Human Genetics* 69:1080-1094.
- Spriggs M (1995) The Lapita Culture and Austronesian prehistory in Oceania. In P Bellwood, J Fox and D Tryon (eds.): *The Austronesians: Historical and Comparative Perspectives*. Canberra: Australian National University.
- Spriggs M (1997) *The Island Melanesians*. Cambridge, MA: Blackwell Publishers.
- Summerhayes G (in preparation) *Island Melanesian Pasts - A view from Archaeology*. In J Friedlaender (ed.): *Population Genetics, Linguistics, and Culture History in the Southwest Pacific*. New York: Oxford University Press.
- Summerhayes G, and Allen J (1993) The transport of Mopir obsidian to Late Pleistocene New Ireland. *Archaeology in Oceania* 28:145-149.
- Wasserman H (1965) Human pigmentation and environmental adaptation. *Archives of Environmental Health* 11:691-694.
- White J (1972) *Ol Tumbuna: Archaeological excavations in the eastern Central Highlands, Papua New Guinea*. Canberra: Research School of Pacific Studies.
- Wurm S (1982) *Papuan Languages of Oceania*. Philadelphia: J Benjamins of North America.

- Yao Y, Lu X, Luo H, Li W, and Zhang Y (2000) Gene admixture in the silk road region of China: evidence from mtDNA and melanocortin 1 receptor polymorphism. *Genes and Genetic Systems* 75:173-178.
- Ye J, Parra E, Sosnoski D, Hiester K, Underhill P, and Shriver M (2002) Melting curve SNP (McSNP) genotyping: a useful approach to SNP genotyping in forensic science. *Journal of Forensic Science* 47:593-600.
- Yeh F, and Boyle T (1997) Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* 129:157.

Chapter 4

A Candidate Gene Study for Signals of Natural Selection Using Pairwise Locus-specific F_{ST}

Introduction

Skin pigmentation is a trait that follows geographic clines largely coincident with levels of ultraviolet radiation (UVR) striking the surface of the earth (Relethford, 1997; Jablonski and Chaplin, 2000). The association of darker pigmentation with high UVR and lighter skin with low UVR has led to the development of a number of different hypotheses that seek to explain normal variation in human skin pigmentation in terms of natural selection (Cowles, 1959; Blum, 1961; Wasserman, 1965; Loomis, 1967; Walter, 1971; Post et al., 1975; Branda and Eaton, 1978; Mackintosh, 2001). Alternatively, sexual selection may have played role in the evolution of pigmentation variation (Darwin, 1871; Diamond, 1992). While it is easy to envision this taking place in small, isolated populations, sexual selection acting in a geographically expanding population might achieve a similar effect across a broader region. Although it is difficult to determine which hypothesis (or which combination of them, as they are not mutually exclusive) best explains the distribution of pigmentation variation observed in the human species, the idea that natural selection plays some sort of role is widely accepted (Relethford, 2000).

An improved understanding of the genes underlying normal variation in human skin pigmentation, as well as advances in molecular genetics that allow for the identification of variation at the DNA sequence level, make it now possible to address two pertinent questions in the evolution of human skin pigmentation variation. The first is whether or not we can detect signals of selection in pigmentation candidate genes

believed to be responsible for differences between groups that are notably different in skin pigmentation (such as Europeans and West Africans). The second question addresses the issue of phenotypic convergence. Specifically, are we able to determine whether or not populations that are similar in pigmentation phenotype are similar due to the same genetic mutations, or if instead these populations have reached adaptive phenotypes due to distinct genetic events? If the latter, this would suggest that a particular pigmentation phenotype, such as light skin, has evolved multiple times over the course of human evolution.

In an effort to address these questions this paper uses a method for detecting signals of selection in the human genome that compares pairwise locus-specific F_{ST} values for pigmentation candidate SNPs to empirical pairwise locus-specific F_{ST} distributions. Estimates of divergence based on F_{ST} or similar measures across the human species range from 5-15% (Wright, 1951; Cavalli-Sforza et al., 1994; Barbujani et al., 1997; Jorde et al., 2000; Long and Kittles, 2003; Tishkoff and Verrelli, 2003; Watkins et al., 2003; Kidd et al., 2004). However, these are values averaged across many markers, and individual markers will have values that fall above or below this mean. We would expect markers associated with greater population differentiation due to directional selection to show high locus-specific F_{ST} values relative to this average.

The concept of using F_{ST} to detect signals of natural selection was first proposed by Lewontin and Krakauer in 1973, and is based on the fact that while demographic processes are expected to affect all regions of the genome equally, natural selection acts in a locus-specific manner (Cavalli-Sforza, 1966). However, a serious drawback to this approach is that these distributions can be affected by demographic processes (Lewontin

and Krakauer, 1975; Nei and Maruyama, 1975; Robertson, 1975; Nei and Chakravarti, 1977). Bowcock et al. (1991) examined patterns of variation across 100 RFLP markers in five diverse human populations and compared them to simulations of F_{ST} distributions under neutrality. While they observed that some markers did show departures from this simulated distribution, the authors acknowledged that variation in population size could also cause these departures from neutrality and that such departures could be difficult to model. Beaumont and Nichols (1996) used a similar analysis based on the expected distribution of F_{ST} under an island model of population structure. However, this method may not be robust in cases where the actual population history of the groups under study involves branching events or unequal admixture between populations (Vitalis et al., 2001).

Recent work by Black et al. (2001) and Akey et al. (2002) suggests that the limitations of Lewontin's method may be overcome when an empirical distribution of F_{ST} values is used instead of a simulated neutral distribution. This is because the empirical distribution will take into account the (unknown) demographic processes affecting the shape of the F_{ST} distribution. Until recently generating such an empirical distribution was difficult due to the large number of markers that would need to be typed. However, with the development of microarray genotyping technologies such as the Affymetrix (Santa Clara, CA) Mapping 10K Array Chip it is now possible to type an individual for a very large number ($n > 10,000$) of SNPs in a single experiment. By typing multiple individuals from different populations it is possible to generate allele frequencies for SNPs on the chip and to use these allele frequencies to calculate the empirical pairwise locus-specific F_{ST} distributions for different human population pairs.

In this paper pairwise locus-specific F_{ST} values for ten pigmentation candidate SNPs from six different pigmentation candidate genes were compared to the empirical locus-specific pairwise F_{ST} distributions for six geographically distinct human populations. The purpose was to determine if any of these SNPs showed a pattern of divergence (high F_{ST}) indicative of directional selection between populations that are different in pigmentation phenotype as well as to identify SNPs with high divergence between populations similar in pigmentation phenotype (suggesting phenotypic convergence). It is important to state that such tests are only indicative of selection acting in the *region* of the candidate SNPs tested, and not necessarily on the SNP itself. It may be that the particular SNPs identified in this work are not in fact the functional variants but are instead linked to (on the same haplotype as) the functional SNP that selection has acted upon.

I have outlined the four possible scenarios for allele frequency and phenotypic similarity in Table 4.1 and briefly describe them here. The first type of comparison involves populations that are different in pigmentation phenotype, such as Europeans or East Asians vs. Island Melanesians or West Africans. It is possible that the candidate SNP will show high locus-specific pairwise F_{ST} between these groups. If this SNP is functionally associated with normal variation in pigmentation (rather than simply showing high allele frequency differences between the two populations), then these high pairwise locus-specific F_{ST} values suggest that the gene (or haplotype) in question may have been subject to directional selection. However, it is also possible that the candidate SNP may show low levels of divergence between the two populations. Such a SNP is unlikely to have been subject to directional selection in shaping pigmentation differences

between the two groups. The second type of comparison involves populations that are similar in skin pigmentation phenotype (e.g. Europeans and East Asians or Island Melanesians and West Africans) and addresses issues of phenotypic convergence. Under this type of comparison we might expect a candidate SNP to show low divergence (low F_{ST}) between the populations if it is affecting pigmentation in both groups in a similar manner. However, if a pigmentation SNP shows high F_{ST} between the two populations, and if it is known to be associated with normal variation in human skin pigmentation then a possible scenario of phenotypic convergence emerges.

Table 4.1: Possible outcomes of pairwise locus-specific F_{ST} comparisons between populations

Different Pigmentation	Similar Pigmentation	Low IsF_{ST}	High IsF_{ST}	Conclusions
X		X		no effect on pigmentation between groups; may affect variation within a group
X			X	if associated with variation between groups → possible target of selection
	X	X		if associated with pigmentation → having similar effect in both groups
	X		X	if associated with pigmentation → possible case of evolutionary convergence

Using this pairwise locus-specific F_{ST} approach on six pigmentation candidate genes we have identified two genes, *OCA2* and *ASIP*, that may have been subject to directional selection between darkly and lightly pigmented populations. We also observe that two of the genes tested, *TYR* and *MATP*, are associated with lighter skin pigmentation in European populations. East Asians have lower frequencies of these

European variants in these genes (leading to high pairwise locus-specific F_{ST} between these two groups), and presumably the haplotype and functional variant associated with it. This suggests that the lighter skin of Europeans and East Asians is due in part to phenotypic convergence. We confirm that these genes are associated with normal variation in pigmentation using admixture mapping, and discuss the evolutionary implications of these results.

Materials and Methods

Samples

Affymetrix Chip

Twenty individuals each from the following populations were typed on the Affymetrix 10K WGA Mapping Array Chip: West African (Mende from Sierra Leone), Island Melanesian (Nasioi from Bougainville), South Asian (Indians from Andhra Pradesh), Native American (Nahua from Mexico), East Asian (Chinese and Japanese), and European (Spanish from Valencia) using the methods described in Shriver et al. (Shriver et al., 2005)

Pigmentation Candidate Gene Typing.

Pigmentation candidate gene SNPs were typed in the same 20 individuals from each population typed on the Affymetrix Chip, as well as on an additional sample

averaging 40 individuals from the same or similar populations. It was possible to sample individuals from the same populations typed on the Affymetrix Chip for the Island Melanians (n = 44; samples were from other neighborhoods on the same island), and Europeans (n = 42). The additional 40 East and South Asian individuals were unadmixed Chinese (n = 46) and South Asian (n = 45) residents of Trinidad and Tobago (collected by Tamiko Brown, UWI). The additional 45 West African individuals were African-American individuals of the Gullah population of South Carolina observed to have 100% West African ancestry by Parra et al. (2001). The additional Native American samples were comprised of 14 Nahua speakers and 33 Mayans (collected by Ken Weiss and Anne Buchanan). The four SNPs that showed signals of selection in the pairwise locus-specific F_{ST} study (*ASIP* A8818G, *OCA2* A355G, *TYR* A192C, and *MATP* C374G) were also typed in the full CEPH Diversity Panel (Cann et al., 2002).

***MATP* Genotype-phenotype Association**

To determine the association of *MATP* with normal variation in skin pigmentation the *MATP* C374G SNP was genotyped in 362 African-American and African-Caribbean individuals for whom quantitative measures of skin pigmentation as well as individual ancestry estimates were available (Shriver et al., 2003). These populations were selected because admixed populations are well-suited for gene mapping due to the long-range linkage disequilibrium (LD) that is generated by admixture.

Marker Selection

Pigmentation candidate SNPs were selected for inclusion in this study because of either allele frequency differences between populations, their location within genes believed to have an effect on pigmentation variation, or because of previously reported associations with normal variation in pigmentation variation. All SNPs are nonsynonymous with the exceptions of *OCA2* A355G, *TYRPI* A209T, and *MC1R* G314A; the *ASIP* A8818G SNP was located in the promoter region.

PCR and Genotyping Methods

Reference SNP numbers and genotyping conditions for the ten SNPs typed in this study can be found in Chapter Three.

F_{ST} Estimation and Percentile Rank Calculation

Unbiased estimates of Weir and Cockerham's F_{ST} were calculated as in Akey et al. (2002). Locus-specific pairwise F_{ST} percentile rank for pigmentation SNPs was determined using the following equation:

$$\text{Rank (x)} = \# \text{ loci} > \text{pairwise locus specific } F_{ST}(x) / \text{total } \# \text{ loci}$$

The higher the pairwise locus-specific F_{ST} value for the pigmentation SNP the more significant its percentile rank. SNPs whose pairwise locus-specific F_{ST} values fell into the top 5% of the empirical distribution were identified as showing significant divergence. This one-tailed test was used because we were only interested in those loci

showing evidence of directional selection; these would fall in the upper tail of the distribution. If we were concerned about both balancing and directional selection, a two-tailed test identifying markers falling in the lower and upper 2.5% of the empirical distribution would have been appropriate.

Phylogenetic Tree Construction

Population trees were constructed using the neighbor-joining method (Nei and Saitou, 1987) as implemented in MEGA 2.1 (Kumar et al., 2001) using average pairwise F_{ST} values from the panel of 11,078 autosomal Affymetrix SNPs as a pairwise distance measure. Locus-specific trees were constructed in the same manner, using the pairwise locus-specific F_{ST} values for pigmentation SNPs considered separately.

Genotype-phenotype Associations at *MATP*

We tested for an association between the *MATP* 374*C variant and pigmentation phenotype in the African-American and African-Caribbean samples using a standard ANOVA with both sex and individual ancestry as covariates. While admixed populations are useful in gene mapping studies because of the allelic association that results as part of the admixture process, failure to control for the population stratification that this process also generates can lead to a number of false-positive associations between the trait of interest and alleles that are merely associated with ancestry. To control for variation in individual ancestry we used maximum likelihood (ML) estimates

(as described in Hanis et al., 1986) of individual West African ancestry as a covariate in our ANOVA tests. These ML estimates were based on allele frequencies at 34 ancestry informative markers (AIMs) described in Shriver et al. (2003). This work demonstrated that this test is conceptually an admixture mapping test, the results of which are highly correlated with results of Paul McKeigue's ADMIXMAP program (<http://www.lshtm.ac.uk/eph/eu/GeneticEpidemiologyGroup.htm>).

Results

Pairwise Locus-specific F_{ST}

Our calculations of the pairwise locus-specific F_{ST} distributions for each of the six populations (15 pairwise combinations) uses the empirical distribution of F_{ST} values, which inherently takes demographic history into account. To illustrate this, the distribution of pairwise locus-specific F_{ST} values for two population pairs, Island Melanesians/ Native Americans and Europeans/South Asians, is shown in Figure 4.1. The distributions of pairwise locus-specific F_{ST} values for each of the population pairs are quite different. While in both cases there is a high proportion of low pairwise locus-specific F_{ST} variants (consistent with a recent common origin of modern humans), it is greater in the European/South Asian comparison. In contrast, the distribution of values for the Island Melanesian/Native American pair shows a longer tail, indicating a greater number of high pairwise locus-specific F_{ST} SNPs. These results demonstrate the importance of being able to take population histories into account using empirical, rather

than theoretically derived or simulated, pairwise locus-specific F_{ST} distributions. Specifically, while a pairwise locus-specific F_{ST} value of 0.157 would be considered significantly high (falling in the top 5% of pairwise locus-specific F_{ST} values) in the European/South Asian distribution, this value would not be significantly high for the Island Melanesian/South Asian distribution, this value would not be significantly high for the Island Melanesian/South Asian distribution (0.396 would be the pairwise locus-specific cut-off value for this population pair).

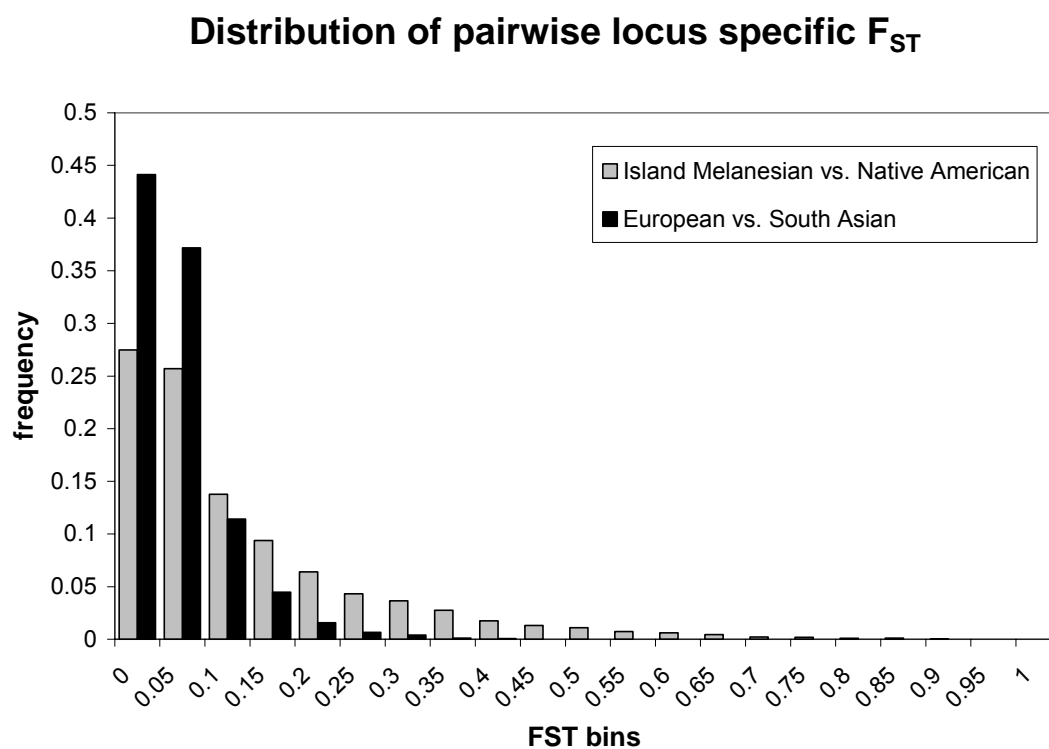


Figure 4.1: Pairwise locus-specific F_{ST} distributions for two population pairs: Island Melanesians vs. Native Americans and Europeans vs. South Asians

Allele frequencies for each of the six populations were calculated for the ten pigmentation candidate SNPs. These are shown in Table 4.2. Mean empirical pairwise F_{ST} values for each possible population pairing and pairwise locus-specific F_{ST} for the

four SNPs that showed signals of selection in at least one population pair are shown in Table 4.3. *OCA2* A355G shows significantly high pairwise locus-specific F_{ST} between West Africans and Europeans. *ASIP* A8818G shows significantly high pairwise locus-specific F_{ST} between West Africans and South Asians, Native Americans, and Europeans. Both *TYR* A192C and *MATP* C374G showed significantly high pairwise locus-specific F_{ST} for all European population pairs. Four of the SNPs considered in this study (*TYR* C308G, *TYR* A402G, *TYRPI* A209T, and *MATP* A272G) showed low variation across all populations (Table 4.2). As such it was not possible to test for signals of selection using these four loci. The remaining two SNPs, *MC1R* A92G and *MC1R* A314G, while polymorphic in our populations, did not exhibit significantly elevated levels of divergence between any population pair.

The mean pairwise F_{ST} values based on the autosomal SNPs in the Affymetrix 10K panel and the pairwise locus-specific F_{ST} values for the four pigmentation SNPs showing signals of divergence were used to draw the neighbor-joining networks shown in Figures 4.2, 4.3, 4.4, 4.5, and 4.6. The high degree of European divergence at *TYR* A192C and *MATP* C374G is evident by the very long branch lengths separating Europeans from all other populations. The high pairwise locus-specific F_{ST} values between Europeans and East Asians at these loci are notable, as they suggest the possibility of independent origins of the light skin phenotype in each population. The population divergence at *OCA2* A355G and *ASIP* A8818G is suggestive of a possible differentiation between lightly and darkly pigmented populations.

Table 4.2: Allele frequencies by population

	n	TYR 192*A	TYR 308*C	TYR 402*A	ASIP 8818*A	TYRP1 209*A	OCA2 355*A	MATP 272*A	MATP 374*C	MC1R 92*G	MC1R 314*G
Island Melanesian (Nasioi)	61	0.00	0.00	0.00	0.51	0.00	0.31	0.00	0.00	0.39	0.29
East Asian (Chinese/Japanese)	61	0.00	0.00	0.00	0.41	0.00	0.47	0.00	0.03	0.44	0.40
South Asian (Andhra Pradesh)	63	0.09	0.00	0.00	0.36	0.00	0.42	0.00	0.13	0.38	0.06
Native American (Nahua)	65	0.02	0.00	0.02	0.05	0.00	0.48	0.00	0.12	0.13	0.08
European (Spanish)	44	0.51	0.05	0.07	0.25	0.00	0.50	0.00	0.25	0.20	0.12
West African (Mende)	66	0.02	0.00	0.03	0.26	0.02	0.07	0.01	0.10	0.51	0.02

Table 4.3: Average pairwise F_{ST} for the Affymetrix 10K WGA chip and locus-specific F_{ST} for four pigmentation candidate SNPs. P-values are shown in parentheses.

	Island Melanesian	East Asian	South Asian	Native American	European
Affymetrix Chip					
East Asian	0.112				
South Asian	0.111	0.067			
Native American	0.157	0.092	0.104		
European	0.141	0.103	0.048	0.123	
West African	0.182	0.157	0.125	0.189	0.137
TYR					
East Asian	0.000 (1.000)	-			
South Asian	0.019 (0.566)	0.019 (0.0436)	-		
Native American	0.000 (1.000)	0.000 (1.000)	0.009 (0.606)	-	
European	0.347 (0.045)	0.347 (0.017)	0.274 (0.004)	0.333 (0.035)	-
West African	0.000 (1.000)	0.000 (1.000)	0.010 (0.646)	0.000 (1.000)	0.333 (0.043)
ASIP					
East Asian	0.037 (0.445)	-			
South Asian	0.061 (0.361)	0.000 (1.000)	-		
Native American	0.269 (0.115)	0.124 (0.147)	0.094 (0.231)	-	
European	0.127 (0.243)	0.024 (0.502)	0.009 (0.436)	0.039 (0.464)	-
West African	0.149 (0.282)	0.323 (0.065)	0.377 (0.023)	0.687 (0.011)	0.496 (0.011)
OCA2					
East Asian	0.037 (0.444)	-			
South Asian	0.011 (0.633)	0.003 (0.580)	-		
Native American	0.042 (0.492)	0.000 (1.000)	0.005 (0.646)	-	
European	0.069 (0.388)	0.001 (0.637)	0.019 (0.345)	0.000 (1.000)	-
West African	0.053 (0.523)	0.167 (0.212)	0.114 (0.235)	0.175 (0.253)	0.348 (0.039)
MATP					
East Asian	0.004 (0.634)	-			
South Asian	0.031 (0.490)	0.013 (0.488)	-		
Native American	0.027 (0.580)	0.010 (0.562)	0.000 (1.000)	-	
European	0.747 (0.001)	0.717 (0.000)	0.624 (0.000)	0.636 (0.003)	-
West African	0.020 (0.545)	0.004 (0.751)	0.000 (1.000)	0.000 (1.000)	0.654 (0.003)

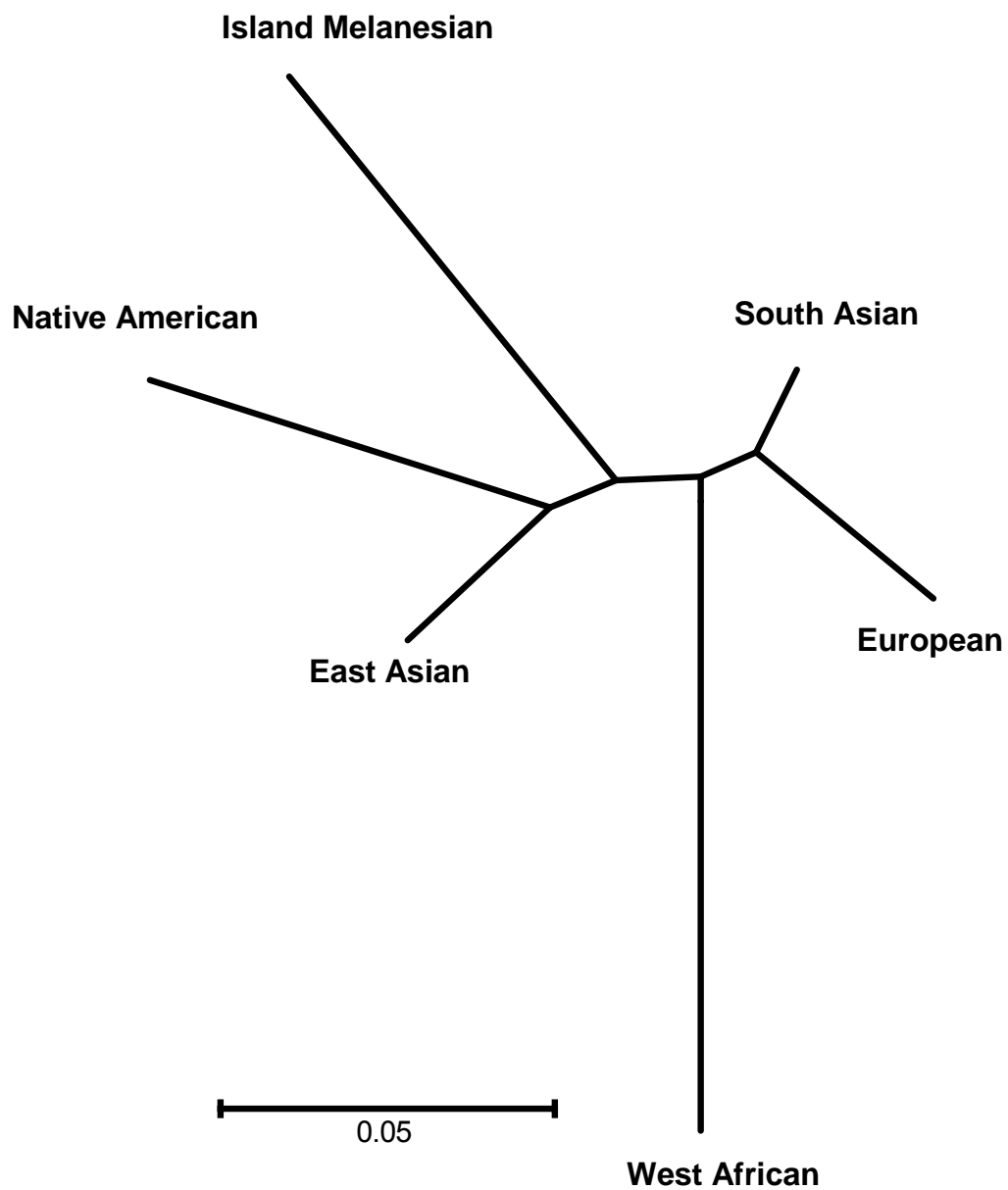


Figure 4.2: Population tree based on average pairwise locus-specific F_{ST} across the 11,078 SNPs typed on the Affymetrix 10K chip

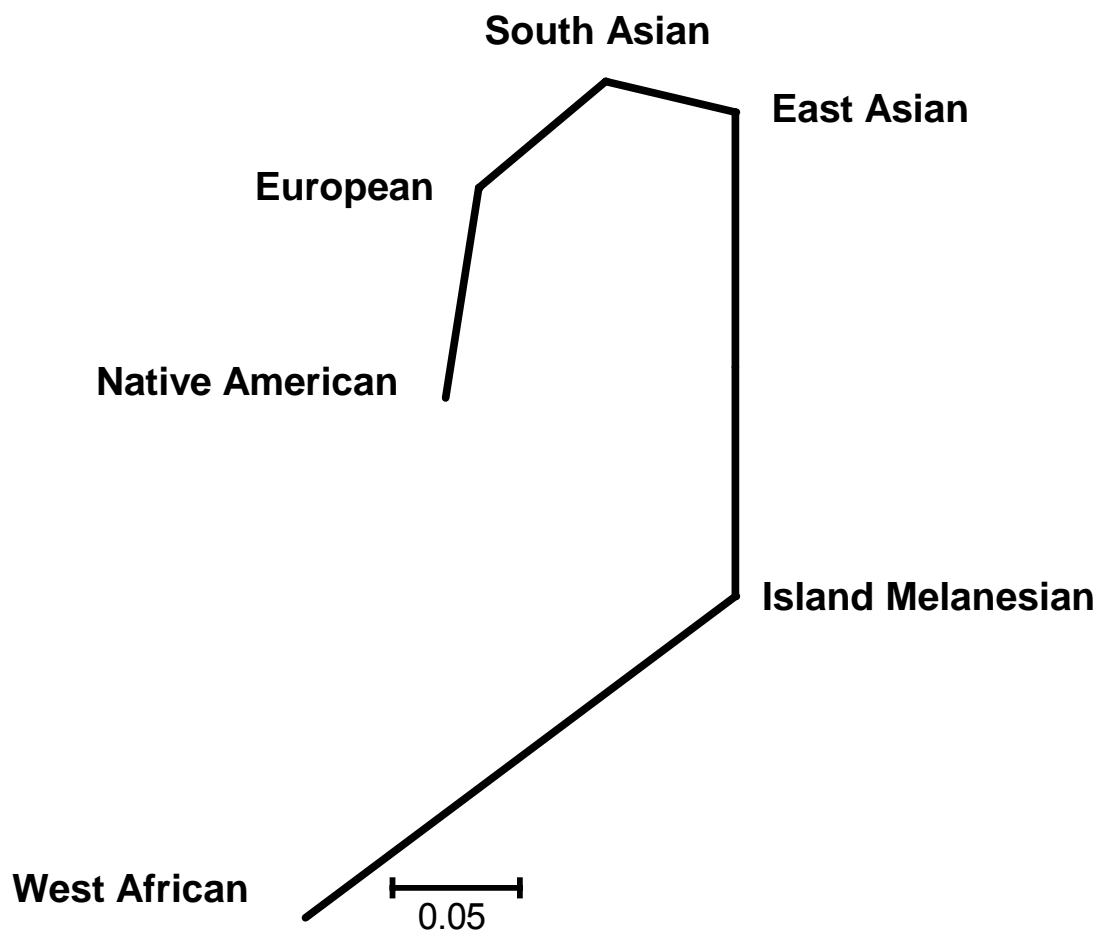


Figure 4.3: Locus-specific tree for *ASIP* A8818G

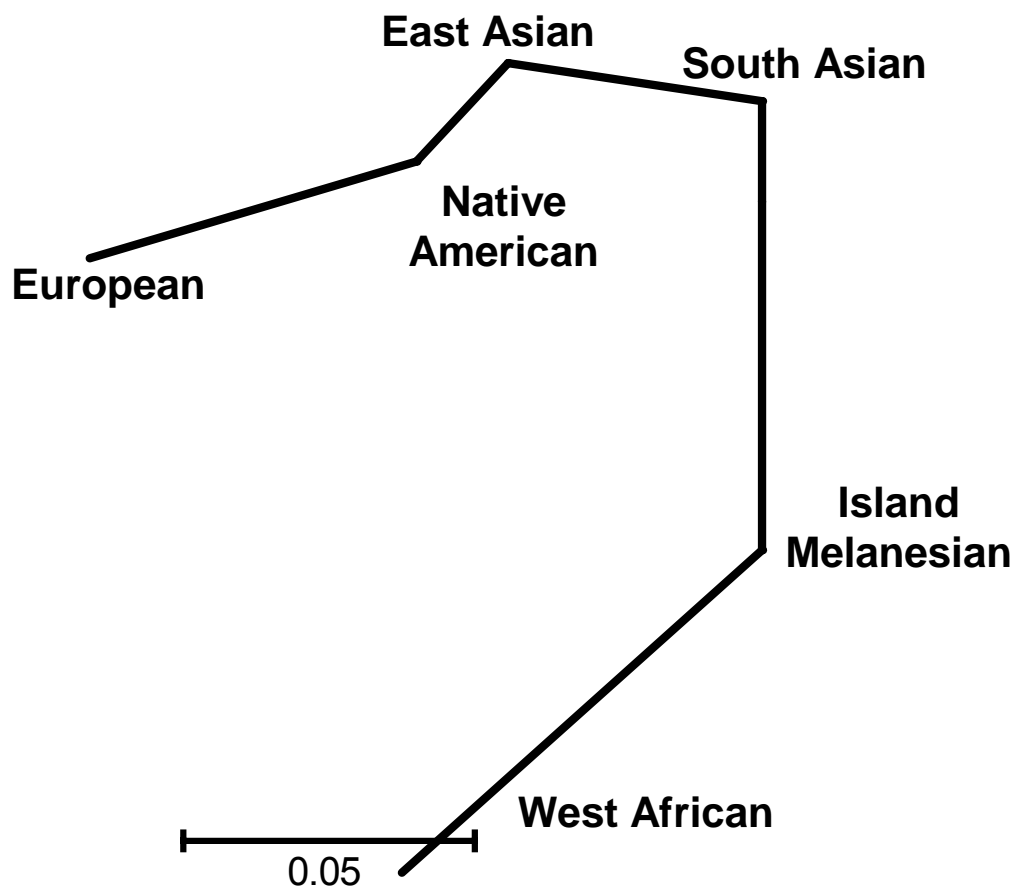


Figure 4.4: Locus-specific tree for *OCA2* A355G

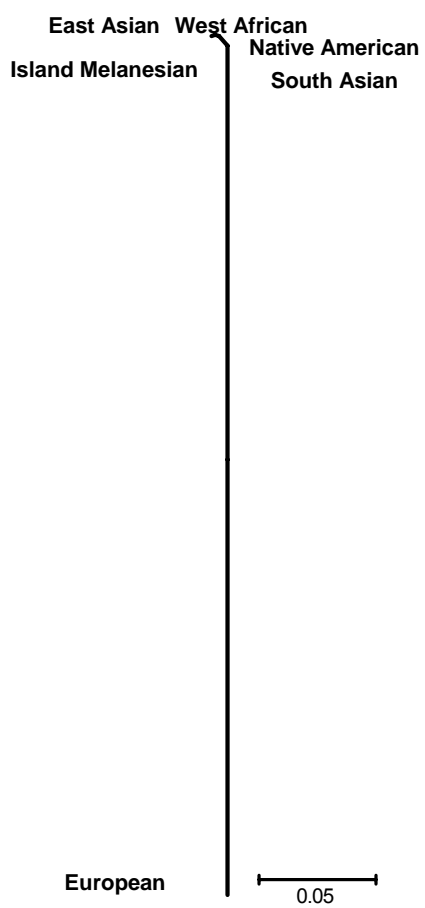


Figure 4.5: Locus-specific tree for *TYR* A192C

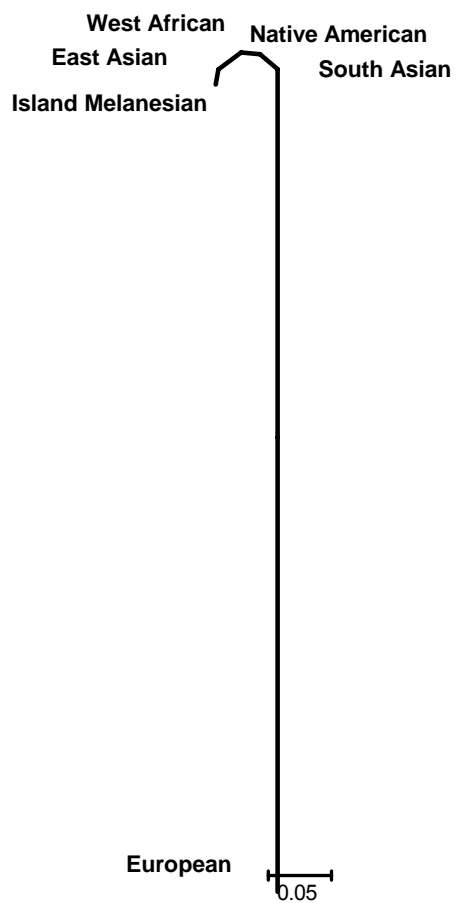


Figure 4.6: Locus-specific tree for *MATP* C374G

Previous work has shown that *TYR* A192C and *OCA2* A355G are associated with normal variation in skin pigmentation in African-American and African-Caribbean populations (Shriver et al., 2003). *ASIP* A8818G is associated with dark hair, eyes, and skin in European populations (Kanetsky et al., 2002) and recently Bonilla et al. (2005) demonstrated linkage using admixture mapping for *ASIP* and normal variation in skin pigmentation in African-Americans. Nakayama et al. (2002) observed high allele frequency differences between populations at *MATP* C374G (Nakayama et al., 2002), and while this work was ongoing Graf et al. (2005) observed that this variant was associated with dark hair and eyes in Europeans.

***MATP* C374G Genotype-phenotype Association**

To determine if the *MATP* C374G SNP was associated with normal variation in skin pigmentation it was typed in a combined sample of 362 African-American and African-Caribbean individuals for whom quantitative measures of skin pigmentation (M index) and individual ancestry estimates were available (Shriver et al., 2003). Mean skin M index values vary across the *MATP* C374G genotypes, with CC homozygotes showing the lowest M index values (42.6) and GG homozygotes showing the highest (57.4). Heterozygotes show an intermediate M index value (51.0). We observe a significant association ($p < 0.0001$) between *MATP* genotype and skin M index using a standard ANOVA with sex and individual ancestry estimates as covariates. Specifically, the greater the number of copies of the C allele (the allele found almost exclusively in Europeans) the lighter an individual's skin pigmentation.

Pigmentation SNP Allele Frequencies in CEPH-Diversity Samples

To further explore the patterns of geographic variation at the four pigmentation SNPs showing signals of selection in using pairwise locus-specific F_{ST} we also typed these SNPs in the CEPH-diversity panel. The distribution of allele frequencies for each of these four SNPs is illustrated in Figures 4.7, 4.8, 4.9, and 4.10. Comparisons of pairwise locus-specific F_{ST} values at *OCA2* A355G and *ASIP* A8818G suggested that natural selection may have led to a divergence between lightly and darkly pigmented populations at these two genes. If dark pigmentation across all human populations is due to the same genetic mutations in both of these genes we would expect to see similar frequencies of the “dark” alleles in CEPH diversity sample populations known to have darker skin pigmentation. At *TYR* A192C and *MATP* C374G the pairwise locus-specific F_{ST} results suggested strong European differentiation at these loci. As such, we would expect to find the “light” alleles almost exclusively in European populations.

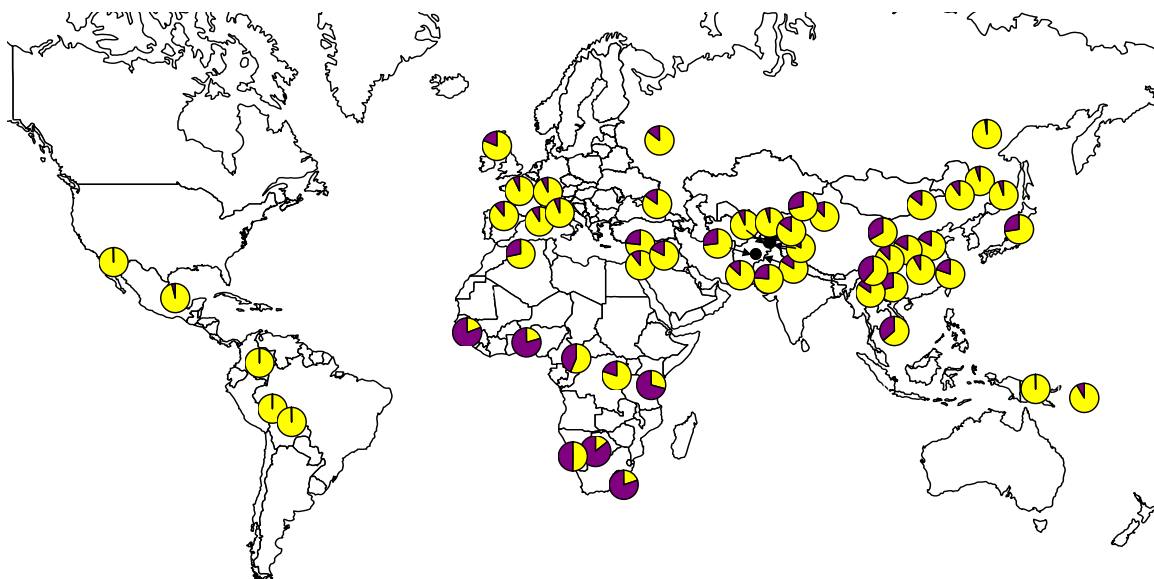


Figure 4.7: Distribution of *ASIP* A8818G in the CEPH-diversity sample set. The “light” allele is shown in yellow.

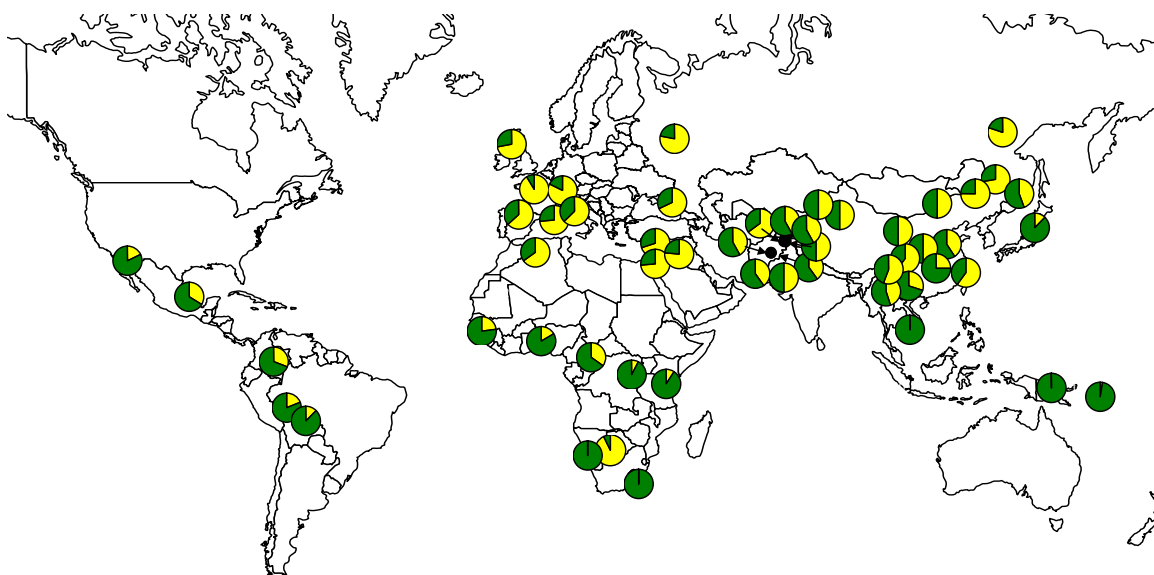


Figure 4.8: Distribution of *OCA2* A355G in the CEPH-diversity sample set. The “light” allele is shown in yellow.

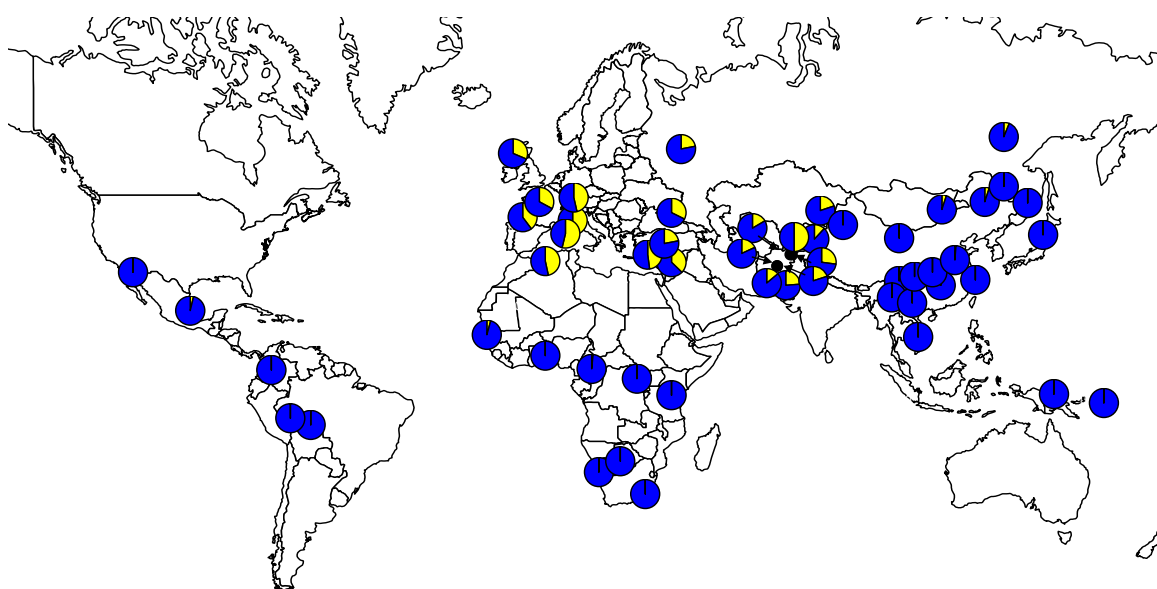


Figure 4.9: Distribution of *TYR* A192C in the CEPH-diversity sample set. The “light” allele is shown in yellow.

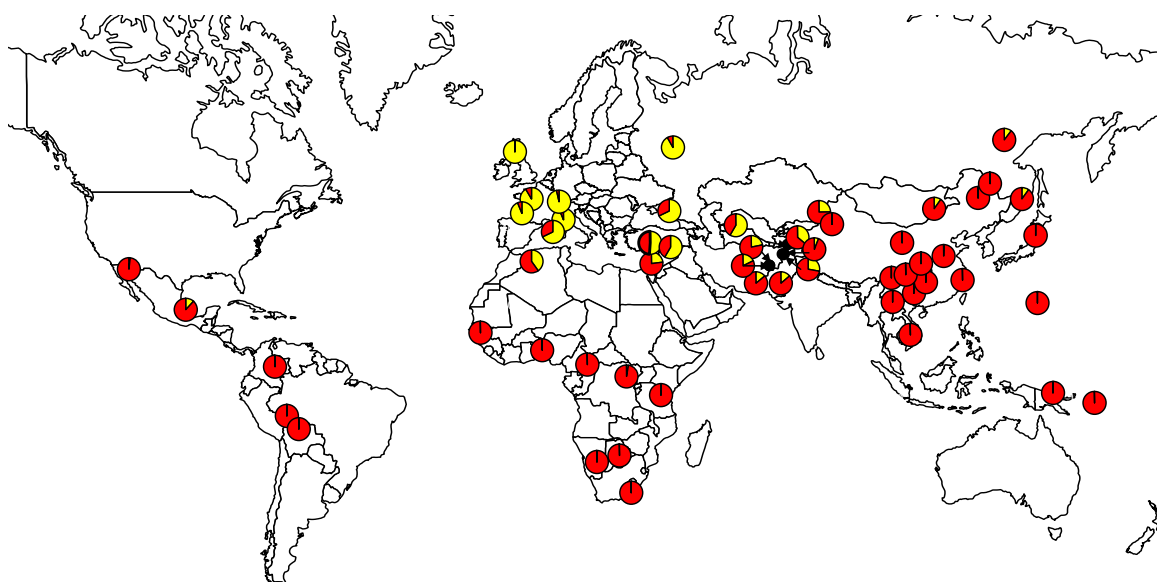


Figure 4.10: Distribution of *MATP* C374G in the CEPH-diversity sample set. The “light” allele is shown in yellow.

At *OCA2*, the 355*A (“light”) allele occurs at high to intermediate frequencies across Europe and Asia, but it also occurs in Native American populations (18-34%) and even among the Bantu-speaking African groups, although at lower frequencies (0-10%). The widespread distribution of this allele suggests that it is not necessarily a new mutation, nor has it been geographically restricted to a single region, although the light allele does occur at lower frequencies across Africa and in Island Melanesia. The elevated frequency of the *OCA2* allele among the lightly-pigmented San of southern Africa is notable.

At *ASIP* the 8818*G (“dark”) allele occurs at its highest frequencies in Africa, although the A (“light”) allele is not uncommon. Interestingly, patterns of *ASIP* variation in the CEPH-diversity panel Island Melanesian samples do not match those observed in our initial pairwise locus-specific F_{ST} study. Specifically, the non-Austronesian (NAN) Melanesian group from Bougainville in the CEPH-diversity panel has the light allele at very high frequency (92%) and the light allele is fixed in the Papuan group from New Guinea. These results are not consistent with those that we observed in a larger sample from Bougainville (see Chapter Three). In that sample of 136 individuals the *ASIP* “light” allele was observed at 47% across our samples from northern Bougainville, although its frequency in the different Bougainville neighborhoods ranged from 32-69%. It was observed at its lowest frequencies among the Papuan speaking Aita (32%). The *ASIP* “light” allele was observed at much higher frequencies elsewhere in Island Melanesia (86% in New Britain, 83% in New Ireland).

The distributions of the “light” alleles at *TYR* A192C and *MATP* C374G are quite different from those observed for *OCA2* A355G and *ASIP* A8818G, and are consistent with predictions made from the pairwise locus-specific F_{ST} results. The *TYR* 192*A allele is primarily restricted to Europe and the Middle East, although it occurs at low frequencies (0-6%) in some East Asian populations. The distribution of the *MATP* 374 C allele is similar, although its frequencies in Europe are even higher than those for *TYR*.

Discussion

The mean pairwise F_{ST} values for the six populations included in this study based on the Affymetrix 10K panel of 11,078 SNPs are consistent with our current understanding of the history of these populations. Those populations that have separated earlier have higher average pairwise F_{ST} values than those populations that have diverged more recently, as illustrated in Figure 1.

OCA2 A355G showed a signal of strong differentiation between West African and European populations, while *ASIP* A8818G showed similar signals between West Africans and Europeans, South Asians, and Native Americans. These results are suggestive of a possible differentiation between darkly pigmented and lightly pigmented populations. Of particular interest is the similarity between West Africans and Island Melanesians, two notably darkly pigmented populations, at these loci. Is it possible that the similar pigmentation phenotype in these groups is due to the same genetic mutations? If so, it suggests that the alleles associated with dark skin might be very old (Rogers et al., 2004), and may have characterized the first modern human migrants out of Africa.

Some believe that original migrants to Australia and Island Melanesia arrived via a coastal route along southern India and southeast Asia (Nei and Roychoudhury, 1993; Cavalli-Sforza et al., 1994; Quintana-Murci et al., 1999; Maca-Meyer et al., 2001; Macaulay et al., 2005). Given that this would keep migrants in a corridor of presumably high UVR, and assuming that the migration occurred relatively rapidly, this may explain the low pairwise locus-specific F_{ST} values between West Africans and Island Melanesians at *ASIP* and *OCA2*. While skin pigmentation is variable in Island Melanesia, the samples from that region used in this study are from Bougainville, an island notable for the very dark pigmentation of its inhabitants (Chapter Two; Norton et al. in press).

The global distribution of the *OCA2* A355G SNP in the CEPH-diversity panel suggests that the “light” allele at this locus is not the result of a recent mutation, as it can be found at intermediate frequencies around the world, with the exception of two populations in Melanesia (“Papuaans” from the New Guinea highlands and Non-Austronesian speakers from Bougainville) and one Southeast Asian population (Cambodian). In Chapter Three this *OCA2* SNP was typed in a larger Island Melanesian sample that includes individuals from the islands of Bougainville, New Britain and New Ireland. That study demonstrated that the light allele does occur at intermediate frequencies (25-64%) across Island Melanesia, although it occurs at its lowest frequency on Bougainville.

The distribution of the *OCA2* light allele across Africa is also interesting as it occurs at intermediate frequencies (0-35%) among Bantu-speaking African populations and at a very high frequency (93%) among the lightly pigmented San. However, the San results should be interpreted with caution as the CEPH-diversity panel contains only

seven San individuals. These results have recently been confirmed among a larger Khoisan sample ($n = 89$), with the light allele occurring at 63 % (Moodley et al. in preparation). Our study suggests that polymorphism at *OCA2* A355G may be very old, and that the increase in frequency of the dark 355*G allele in Africa and Melanesian occurred later in human evolution (subsequent to migration out of Africa). Khoisan speaking groups cluster at the base of human mtDNA (Chen et al., 2000) and Y chromosome phylogenies (Semino et al., 2002), which may lend support to our hypothesis of very old polymorphism at *OCA2* A355G.

The distribution of the *ASIP* “light” and “dark” alleles across the CEPH-diversity sample set is somewhat puzzling. Similar to *OCA2*, we observed that both the light and dark alleles have a widespread global distribution, although some geographic patterns can be discerned. The *ASIP* “dark” allele occurs at higher frequencies (20-86%) in most African populations (including the lightly pigmented San), but is nearly absent in the New World. Surprisingly, we observed the dark allele at very low frequencies in the CEPH-diversity panel Island Melanesian samples despite the initial pairwise locus-specific F_{ST} results that suggested similarities between Island Melanesians and West Africans at *ASIP*. We have previously observed the *ASIP* “dark” allele to occur at intermediate frequencies across Island Melanesia, with elevated frequencies in Bougainville. Thus, the low frequency of it in the NAN Melanesian samples from Bougainville is surprising, despite the small sample size ($n = 21$) of the CEPH-diversity sample. One possible explanation lies in the fact that Island Melanesia is composed of a large number of genetically heterogeneous populations. Early arrivals to the region may

have been subject to a greater degree of isolation than is observed today, or even than what existed ~3,200 years ago with the arrivals of Austronesian speaking groups.

Similarity between Africans and Island Melanesians at *OCA2* and *ASIP* would suggest that the dark skin of both groups is due in part, although not exclusively, to the same genetic causes. This would be consistent with hypotheses predicting an early southern migration out of Africa and into Southeast Asia and Island Melanesia. While the pairwise locus-specific F_{ST} results suggested this as a possible scenario, the patterns of variation in the CEPH-diversity panel are not consistent with this hypothesis at *ASIP*. This could be due to the fact that Island Melanesians are a heterogeneous group to begin with, and so it is possible that the samples included on the CEPH-diversity panel simply represent a group that does not share these mutations. These results reflect the polygenic nature of pigmentation phenotype. While some Island Melanesians (perhaps those descended from the earliest migrants to the region) may share the same pigmentation variants with Africans, it may be that others can attribute their dark pigmentation levels to different genetic mutations (perhaps common in some Southeast Asian populations). Given the complex population history of the region this is not surprising. To gain a better understanding of similarities in pigmentation genes between darkly-pigmented populations it would be extremely useful to be able to type samples from dark-skinned tribal groups from South Asia or Australian aboriginal populations.

Both *TYR* A192C and *MATP* C374G showed pairwise locus-specific F_{ST} values that provided strong evidence for European-specific divergence at these loci. This divergence is notable because it separates Europeans from populations that are very different in skin pigmentation phenotype (e.g. West Africans and Island Melanesians) as

well as from a population that is more similar in pigmentation phenotype (East Asians). The sharp differences between Europeans and East Asians at these loci suggest a possible case of phenotypic convergence, in which the lighter skin pigmentation of both groups arose through independent means.

The distribution of the “light” alleles at both *TYR* A192C and *MATP* C374G across the CEPH-diversity panel is consistent with the hypothesis that both of these polymorphisms arose relatively recently and occur at high frequencies in European populations. Genotyping these loci in the CEPH-diversity panel did demonstrate the light alleles at both loci occur at declining frequencies as one moves away from northern Europe and into western Asia. For both genes the light allele occurs at its highest frequencies in Europe and decreases in frequency as one moves west across the Middle East and into Central Asia. This distribution pattern and the confirmation that both loci are linked to normal pigmentation variation in African-Americans and African-Caribbeans, strengthen the case for phenotypic convergence. We do not see the light alleles occurring in the East Asian populations, or in the lightly pigmented San. If these alleles are specifically associated with lighter skin in European populations then we would expect the increase in frequency of these alleles to have occurred relatively recently, following modern human expansion into Europe (~ 30-35 KYA).

It is important to discuss briefly two loci that did not provide a signal of selection in our genome scan: *MC1R* G92A and *MC1R* G314A. *MC1R* is one of the most well-studied pigmentation candidate genes, largely due to its clear effects on normal variation in human skin and hair pigmentation. Various *MC1R* mutations are associated with a red hair, fair skin, and freckled phenotype (Valverde et al., 1995; Box et al., 1997; Smith et

al., 1998; Flanagan et al., 2000; Bastiaens et al., 2001). The high frequencies of some of these variants in European populations have led to speculation that *MC1R* may have been under positive selection, in keeping with predictions made by the Vitamin D hypothesis. Sequence-based tests of natural selection in the coding region of *MC1R* have drawn conflicting conclusions (Rana et al., 1999; Harding et al., 2000), while a study of the promoter region raised the possibility of purifying selection or population-size reduction in some regions of the promoter in African populations but could not distinguish between diversifying selection or a relaxation of functional constraint among European and East Asian populations (Makova et al., 2001). Our pairwise locus-specific F_{ST} based genome scan did not identify *MC1R* as having been subject to directional selection in the populations that we examined. This may in part be due to the loci that were selected for study. One, *MC1R* G314A, is a synonymous SNP that is polymorphic across African, European, and East Asian populations. As it is a synonymous variant we might not expect to observe a strong signal of selection associated with it. The second *MC1R* SNP used in this study, G92A, was originally included as part of the panel of pigmentation SNPs typed in Island Melanesians (Chapter Three). It was selected for use in that study because of its reported polymorphism in East Asian populations. It is possible that had we examined pairwise locus-specific F_{ST} values for strong red hair color (RHC) alleles in *MC1R* we may have identified a signal of selection at this gene.

Conclusions

This work used a measure of population divergence, pairwise locus-specific F_{ST} , to identify pigmentation genes that may have been subject to natural selection. While natural selection has long been hypothesized to play a role in normal pigmentation variation, intensive studies of sequence variation in pigmentation candidate genes have been primarily limited to *MC1R*. The use of pairwise locus-specific F_{ST} to detect signals of selection makes it possible to rapidly screen multiple pigmentation candidate genes rather than focusing on a single gene or a single region within a gene. By comparing pairwise locus-specific F_{ST} values for different pigmentation candidate SNPs to empirical distributions of pairwise locus-specific F_{ST} values typed in the same or similar populations it was possible to take the demographic histories of these populations into account, strengthening the case that those loci showing high pairwise locus-specific F_{ST} values have been influenced by natural selection rather than by demographic factors.

Many natural selection-based hypotheses regarding the evolution of human skin pigmentation variation focus on either the importance of dark skin in environments of high UVR (e.g. photoprotection, folic acid hypotheses) or the importance of lighter skin in low UVR (e.g. Vitamin D hypothesis). Our comparison of pairwise locus-specific F_{ST} values of pigmentation SNPs to empirical F_{ST} distributions identified four genes that may have been involved in adaptation to environments of varying UVR. Two of these (*OCA2* A355G and *ASIP* A8818G) show signals of differentiation between darkly pigmented and lightly pigmented populations, and two (*TYR* A192C and *MATP* C374G) show evidence of strong selection in a lightly pigmented population.

West Africans and Island Melanesians, both groups with relatively dark skin pigmentation, showed pairwise locus-specific F_{ST} values that were below the mean pairwise F_{ST} for these two groups at *OCA2* A355G and *ASIP* A8818G. We took this as evidence that the similarities in pigmentation phenotype in these groups might be due to similarities in pigmentation genotype. This would suggest that the dark skin is an ancestral characteristic. This scenario fits well with the idea that at least part of Island Melanesian was settled very early (~ 65kya) by dark-skinned migrants following a southern coastal route out of Africa.

However, when we typed our *OCA2* and *ASIP* markers in the CEPH-diversity panel we observed that the story might not be quite that simple. First, we saw that the *OCA2* light and dark alleles had a wide global distribution, with the frequencies of the *OCA2* light allele in African Bantu groups ranging from 0 – 17%, but occurring in the San at 93%. The high frequency of the light allele in the San (later confirmed in a larger sample) suggests that the *OCA2* “light” allele is an old one. Although it was not observed in the CEPH-diversity panel Island Melanesian groups, the *OCA2* light allele was observed in the larger Island Melanesian samples typed in Chapter Three at frequencies ranging from 64% in New Britain to 25% in Bougainville.

ASIP appears to be an even more complicated example, as the dark allele was observed at very low frequencies among the CEPH-diversity panel Island Melanesians, although it did occur at elevated frequencies among the larger Island Melanesian sample typed in Chapter Three. The discrepancies between our large Island Melanesian sample and the CEPH-diversity sample may be due in part to the fact that Island Melanesia has had an extremely complex population history characterized by periods of isolation.

While some of the early migrants to the region may share pigmentation gene characteristics with West Africans, other populations may not, perhaps due to drift or due to admixture with later migrants to the region. The smaller sample size of the CEPH-diversity population samples may also be a factor here. It is also worthwhile to consider that the effect of *ASIP* on darker pigmentation may not be as strong as our original pairwise locus-specific results suggested.

Both *TYR* A192C and *MATP* C374G show extremely high pairwise locus-specific F_{ST} values for all possible European population pairings. This suggested strong European divergence at these loci. Additional typing in the CEPH-diversity panel confirmed this, with the light alleles at both these genes almost entirely restricted to European and Middle Eastern populations. This suggests that the similarity in pigmentation between East Asians and Europeans is the result of phenotypic convergence due to different genetic mutations, consistent with predictions made by the Vitamin D Hypothesis. It is hoped that future work will identify pigmentation alleles showing high East Asian divergence to complement those observed for Europeans.

It is important to emphasize that the pairwise locus-specific F_{ST} method used in this study identified markers that warrant further examination in terms of natural selection, not markers that we can say unequivocally have been subject to natural selection. To confirm the signals observed here it will be necessary to examine genetic variation around these SNPs more closely, either by direct sequencing or by constructing SNP or microsatellite haplotypes in the region. Still, we believe that this method has proved its usefulness, as it allowed us to narrow our search from six genes to four. In addition, it identified a gene showing a very strong signal, *MATP*, that may not have been

perceived as a likely candidate for selection at the start of this study, due to the relative lack of information about it.

Literature Cited

- Akey J, Zhang G, Zhang K, Jin L, and Shriver M (2002) Interrogating a high-density SNP map for signatures of natural selection. *Genome Research* 12:1805-1814.
- Barbujani G, Magangni A, Minch E, and Cavalli-Sforza L (1997) An apportionment of human genetic diversity. *Proceedings of the National Academy of Sciences* 94:4516-4519.
- Bastiaens M, ter Hurme J, Gruis N, Bergman W, Westendorp R, Vermeer B, and Bouwes Bavinck J (2001) The melanocortin-1 receptor gene is the major freckle gene. *Human Molecular Genetics* 10:1701-1708.
- Beaumont M, and Nichols R (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London, B* 263:1619-1626.
- Black W, Baer C, Antolin M, and DuTeau N (2001) Population Genomics: Genome-wide sampling of insect populations. *Annual Review of Entomology* 46:441-469.
- Blum H (1961) Does the melanin pigment of human skin have adaptive value? *Quarterly Review of Biology* 36:50-63.
- Bonilla C, Boxill L, Donald S, Williams T, Sylvester N, Parra E, Dios S, Norton H, Shriver M, and Kittles R (2005) The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Human Genetics* 116:402-406.

- Bowcock A, Kidd J, Mountain J, Hebert J, Carotenuto L, Kidd K, and Cavalli-Sforza L (1991) Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. *Proceedings of the National Academy of Sciences* 88.
- Box N, Wyeth J, O'Gorman L, Martin N, and Sturm R (1997) Characterization of melanoocyte stimulating hormone receptor variant alleles in twins with red hair. *Human Molecular Genetics* 6:1891-1897.
- Branda R, and Eaton J (1978) Skin color and nutrient photolysis: an evolutionary hypothesis. *Science* 201:625-626.
- Cann H, de Toma C, Cazes L, Legrand M, Morel V, Piouffre L, Bodmer J, Bodmer W, Bonne-Tamir B, Cambon-Thomsen A, Chen Z, Chu J, Carcassi C, Contu L, Du R, Excoffier L, Ferrara G, Friedlaender J, Groot H, Gurwitz D, Jenkins T, Herrera R, Huang X, Kidd J, Kidd K, Langaney A, Lin A, Mehd iS, Parham P, Piazza A, Pistillo M, Qian Y, Shu Q, Xu J, Zhu S, Weber J, Greely H, Feldman M, Thomas G, Dausset J, and Cavalli-Sforza L (2002) A Human Genome Diversity Cell Line Panel. *Science* 296:261-262.
- Cavalli-Sforza L (1966) Population structure and human evolution. *Proceedings of the Royal Society of London, B* 164:362-379.
- Cavalli-Sforza L, Menozzi P, and Piazza A (1994) *The History and Geography of Human Genes*. Princeton: Princeton University Press.
- Chen Y, Olckers A, Schurr T, Kogelnik A, Huoponen K, and Wallace D (2000) mtDNA variation in the South African Kung and Khwe-and their genetic relationships to

other African populations. *The American Journal of Human Genetics* 66:1362-1383.

Cowles R (1959) Some ecological factors bearing on the origin and evolution of pigment in the human skin. *American Naturalist* 93:283-293.

Darwin C (1871) *The Descent of Man, and Selection in Relation to Sex*. London: J. Murray.

Diamond J (1992) *The Third Chimpanzee*. New York: Harper Collins Publishers.

Flanagan N, Healy E, Ray A, Philips S, Todd C, Jackson I, Birch-Machin M, and Rees J (2000) Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation. *Human Molecular Genetics* 9:2531-2537.

Graf J, Hodgson R, and van Daal A (2005) Single nucleotide polymorphisms in the *MATP* gene are associated with normal human pigmentation variation. *Human Mutation* 25:278-284.

Hanis C, Chakraborty R, Ferrell R, and Schull W (1986) Individual admixture estimates: disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas. *American Journal of Physical Anthropology* 70:433-441.

Harding R, Healy E, Ray A, Ellis N, Flanagan N, Todd C, Dixon C, Sajantila A, Jackson I, Birch-Machin M, and Rees J (2000) Evidence for variable selective pressures at MC1R. *American Journal of Human Genetics* 66:1351-1361.

Jablonski N, and Chaplin G (2000) The evolution of human skin coloration. *Human Evolution* 39:57-106.

- Jorde L, Watkins W, Bamshad M, Dixon M, Ricker C, Seielstad M, and Batzer M (2000) The distribution of human genetic diversity: a comparison of mitochondrial, autosomal, and Y-chromosome data. *The American Journal of Human Genetics* 66:979-988.
- Kanetsky P, Swoyer J, Panossian S, Holmes R, Guerry D, and Rebbeck T (2002) A polymorphism in the agouti signaling protein gene is associated with human pigmentation. *American Journal of Human Genetics* 70:770-775.
- Kidd K, Pakstis A, Speed W, and Kidd J (2004) Understanding human DNA sequence variation. *Journal of Heredity* 95.
- Kumar S, Tamura K, Jakobsen I, and Nei M (2001) MEGA2: Molecular Evolutionary Genetics Analysis software, *Bioinformatics* (submitted). *Bioinformatics* 17:1244-1245.
- Lewontin R, and Krakauer J (1975) Letters to the editors: Testing the heterogeneity of F values. *Genetics* 80:397-398.
- Long J, and Kittles R (2003) Human genetic diversity and the nonexistence of biological races. *Human Biology* 75:449-471.
- Loomis W (1967) Skin pigment regulation of vitamin-D biosynthesis in man. *Science* 157:501-506.
- Maca-Meyer N, Gonzalez A, Larruga J, Flores C, and Cabrera V (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genetics* 2.
- Macaulay V, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F, Taha A, Kassim Shaari N, Maripa Raja J, Ismail P, Zainuddin Z, Goodwin W, Bulbeck D, Bandelth H, Oppenheimer S, Torroni A,

- and Richards M (2005) Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 308.
- Mackintosh J (2001) The antimicrobial properties of melanocytes, melanosomes, and melanin and the evolution of black skin. *Journal of Theoretical Biology* 211:101-113.
- Makova K, Ramsay M, Jenkins T, and Li W (2001) Human DNA sequence variation in a 6.6-kb region containing the melanocortin 1 receptor promoter. *Genetics* 158:1253-1268.
- Nakayama K, Fukamachi S, Kimura H, Koda Y, Soemantri A, and Ishida T (2002) Distinctive distribution of AIM1 polymorphism among major human populations with different skin color. *Journal of Human Genetics* 47:92-94.
- Nei M, and Chakravarti A (1977) Drift variances of FST and GST statistics obtained from a finite number of isolated populations. *Theoretical Population Biology* 11:307-325.
- Nei M, and Maruyama T (1975) Letters to the editors: Lewontin-Krakauer test for neutral genes. *Genetics* 80:395.
- Nei M, and Roychoudhury A (1993) Evolutionary relationships of human populations on a global scale. *Molecular Biology and Evolution* 10:927-943.
- Nei M, and Saitou N (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.
- Post P, Daniels F, and Binford R (1975) Cold injury and the evolution of 'white' skin. *Human Biology* 39:131-143.

- Quintana-Murci L, Semino O, Bandelt H, Passarino G, McElreavey K, and Santachiara-Benerecetti A (1999) Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nature Genetics* 23:437-441.
- Rana B, Hewett-Emmett D, Jin L, Chang B, Sambuughin N, Lin M, Watkins S, Bamshad M, Jorde L, Ramsay M, Jenkins T, and Li W (1999) High polymorphism at the human melanocortin 1 receptor locus. *Genetics* 151:1547-1557.
- Relethford J (1997) Hemispheric difference in human skin color. *American Journal of Physical Anthropology* 104:449-457.
- Relethford J (2000) Human skin color diversity is highest in sub-Saharan African populations. *Human Biology* 75:773-780.
- Robertson A (1975) Gene frequency distributions as a test of selective neutrality. *Genetics* 80:775-785.
- Rogers A, Iltis D, and Wooding S (2004) Genetic variation at the MC1R locus and time since loss of human body hair. *Current Anthropology* 45:105-108.
- Semino O, Santachiara-Benerecetti A, Falaschi F, Cavalli-Sforza L, and Underhill P (2002) Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *American Journal of Human Genetics* 70:265-268.
- Shriver M, Mei P, Parra E, Sonpar V, Halder I, Tishkoff S, Schurr T, Zhadanov S, Osipova L, Brutsaert T, Friedlaender J, Jorde L, Watkins W, Bamshad M, Gutierrez G, Loi H, Matsuzaki H, Kittles R, Argyropoulos G, Fernandez J, Akey J, and Jones K (2005) Large-scale SNP analysis reveals clustered and continuous human genetic variation. *Human Genomics* 2.

- Shriver M, Parra E, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Human Genetics* 112:387-399.
- Smith R, Healy E, Siddiqui S, Falanagan N, Steijlen P, Rosdahl I, and Japques J (1998) Melanocortin 1 receptor variants in Irish populations. *Journal of Investigative Dermatology* 111:119-122.
- Tishkoff S, and Verrelli B (2003) Patterns of human genetic diversity: implicatins for human evolutionary history and disease. *Annual Review of Genomics and Human Genetics* 4:293-340.
- Valverde P, Healy E, Jackson I, Rees J, and Thody A (1995) Variants of the melanocyte stimulating hormone receptor gene are associated with red hair and fair skin. *Nature Genetics* 11:328-330.
- Vitalis R, Dawson K, and Boursot P (2001) Interpretation of variation across marker loci as evidence of selection. *Genetics* 158:1811-1823.
- Walter H (1971) Remarks on the environmental adaptation of man. *Humangenetik* 13.
- Wasserman H (1965) Human pigmentation and environmental adaptation. *Archives of Environmental Health* 11:691-694.
- Watkins W, Rogers A, Ostler C, Wooding S, Bamshad M, Brassington A, Carroll M, Nguyen S, Walker J, Prasad B, Reddy P, Das P, Batzer M, and Jorde L (2003) Genetic variation among world populations: inferences from 100 Alu insertion polymorphisms. *Genome Research* 13:1607-1618.
- Wright S (1951) The genetical structure of populations. *Annals of Eugenics* 15:323-354.

Chapter 5

Conclusions and Future Directions

Introduction

This dissertation has explored phenotypic, genotypic, and evolutionary aspects of human pigmentation variation. This final chapter will discuss issues that arose during the course of this work and directions for future studies. In a phenotypic study in Island Melanesia I observed that pigmentation can vary extensively even within a geographically restricted region. This suggests that forces such as genetic drift and sexual selection can affect pigmentary phenotypes even in the face of strong natural selective pressure. My study of candidate gene variation in Island Melanesia identified possible genotype-phenotype associations at *ASIP*, *OCA2*, and *MC1R*, but population stratification in the region confounds interpretation of these associations. This indicated that an effective study design for mapping pigmentation genes in Island Melanesia should include markers and methods that would control for population substructure. Using pairwise locus-specific F_{ST} I identified four pigmentation candidate genes that may have been influenced by natural or sexual selection. Two of these, *OCA2* and *ASIP*, show signals of divergence between darkly and lightly pigmented populations. The other two genes, *MATP* and *TYR*, show strong signals of European-specific divergence. These studies have enhanced current understanding of phenotypic and genotypic variation in pigmentation candidate genes and the evolutionary forces that may have shaped this variation within and between populations. Over the course of this work methods for improving future studies were identified—these are briefly discussed below.

Ascertainment Bias in SNP Selection

The ten SNPs that were selected for genotyping in Chapters Three and Four were chosen based on reported variation in allele frequencies from the literature or from dbSNP. One potential problem in relying on allele frequencies cited in dbSNP is that many times the SNPs found in this database have been ascertained in small samples enriched for European ancestry. As a result of this ascertainment bias some SNPs that are polymorphic outside of European populations may not be included. Also, the SNPs that are reported to be polymorphic in European populations may not show similar levels of variation outside of Europe. This may be a problem when attempting to detect SNPs associated with pigmentation variation. This was particularly evident in Chapter Three. Six of the ten pigmentation candidate SNPs tested in Chapter Three showed little to no variation across the Island Melanesian sample. While two of these (*TYR* A192C and *MATP* C374G) were polymorphic in European samples the remaining four showed little to no variation across Island Melanesia or in any of the comparison samples of Europeans, South Asians, West Africans, East Asians, and Native Americans. As six of the SNPs selected for genotyping result in amino acid substitutions I expected to see variation in allele frequencies between populations that were different in skin pigmentation, but this was not the case.

For genotype-phenotype association studies in Island Melanesia a more appropriate method to discover SNPs with useful levels of polymorphism would be to sequence short regions of candidate genes in samples of Island Melanesians.

Alternatively, although Island Melanesians are not included in the HapMap sample set

(www.hapmap.org), selecting SNPs typed among the HapMap samples would at least allow for the identification of SNPs that are polymorphic in East Asian, West African and European populations.

A second major issue stemming from ascertainment bias in this dissertation is that four of the SNPs considered here (*TYR* A192C, *OCA2* A355G, *ASIP* A8818G, and *MC1R* G314A) had been previously identified as ancestry informative markers (AIMs) in admixture mapping studies (Parra et al., 1998; Parra et al., 2001; Shriver et al., 2003; Bonilla et al., 2005). A fifth SNP, *MATP* C374G, while not identified as an AIM, was selected because of high allele frequency differences between populations (a defining characteristic of AIMs). In Chapter Four I used pairwise locus-specific F_{ST} to detect signals of natural selection. By including AIMs in the marker panel that I genotyped I increased the chance of identifying significantly high F_{ST} markers because AIMs are by definition high F_{ST} markers. However, since the genes signaled by these AIMs show linkage to skin pigmentation by admixture mapping we can be confident that the high pairwise locus-specific F_{ST} observed at these markers is not an artifact. Although ascertainment bias may have influenced the inclusion of these SNPs in this study, this work has contributed to current knowledge about the geographic distribution of these alleles outside of the West African, Native American, and European populations commonly used in admixture mapping studies.

Identification of SNPs vs. Haplotypes Associated with Pigmentation Variation

Chapters Three and Four examined variation at several SNPs in pigmentation candidate genes. It is important to understand that admixture mapping linkages between phenotype and genotype at a particular SNP like that observed for *MATP* C374G in Chapter Four are indications of linkage between phenotype variation and the gene in question, not necessarily the specific SNP used in that study. This is because linkage disequilibrium generated by admixture creates long blocks of LD—one of the advantages of using admixed populations in gene mapping studies. Since strong, recent selection may generate extended blocks of linkage disequilibrium (LD) around the selected site (although these blocks are not as long as those generated through admixture) it is possible that the pairwise locus-specific F_{ST} tests conducted in Chapter Four may be identifying signals from a functional SNP that occurs some distance from the marker SNP.

While the linkage between marker SNPs typed in this study and pigmentation phenotype may not have identified the exact functional variant associated with pigmentation variation, by typing additional markers or sequencing in the region around the marker SNP it should be possible to identify the relevant functional variants. For example, by typing additional markers or sequencing in the region around the *MATP* C374G SNP it will be possible to construct *MATP* haplotypes. We expect to observe different *MATP* haplotypes between Europeans and other populations, and that the functional variant may occur uniquely on some subset of these haplotypes. Candidates for the functional SNPs on these haplotypes identified by sequencing can then be cloned and tested for a functional effect *in vitro*.

Changes in Pigmentation Over the Course of Human Evolution

Studies of pigmentation variation and variation in pigmentation candidate genes in non-human primates have been primarily restricted to coat color variation (Mundy and Kelly, 2003), but an exploration of variation in skin pigmentation phenotype among non-human primates, particularly in chimpanzees, may be helpful in determining what was the ancestral human skin pigmentation phenotype. In an effort to assess the visibility of different primates to each other as well as to avian predators Sumner and Mollon (2003) made quantitative measurements of skin and coat color pigmentation in different species of primates. Unfortunately their measurements of skin pigmentation were primarily restricted to the brightly pigmented male mandrills (Sumner and Mollon, 2003) and do not include chimpanzees or bonobos, which would be a more appropriate comparison to make when trying to determine the ancestral color of human skin.

There is variation across great ape species in terms of skin pigmentation. Some apes, such as gorillas, have darker skin on their faces and beneath their fur, while others like chimpanzees have skin that is lighter in color. Again, there have been few studies of how (or if) skin pigmentation among the great apes changes with age, but it is known that the infants of some species, such as the Borneo orangutan [*Pongo pygmaeus*, (Rowe, 1996)] and the common chimpanzee [*Pan troglodytes*, (Post et al., 1975)] are born with pink faces that darken with age. While adult chimpanzees are more darkly pigmented than infants (especially on the face and hands), there is variation across adult individuals. As chimpanzees and bonobos represent our closest living relatives it is interesting to

consider if human ancestral pigmentation was relatively light, as in modern chimpanzees, or if it was darker (perhaps more similar to gorillas).

It is important in this discussion to define what is meant by “ancestral”. For example, does “ancestral” human skin pigmentation refer to the phenotype of a very early hominid such as *Sahelanthropus tchadensis* or is it more appropriate to think of a more recent ancestor like *Homo ergaster*? For the discussion that follows I am going to consider the question of ancestral human pigmentation in reference to the phenotype of the first anatomically modern humans (AMH). The oldest evidence for AMH comes from the site of Herto, Middle Awash, Ethiopia, and is dated to 160-154 KYA (White et al., 2003). As such, the following discussion will consider the skin pigmentation phenotype of AMH prior to its expansion out of Africa.

Current technical limitations prevent the recovery of DNA from bones much older than 100 KYA (Paabo and Wilson, 1991), and as most ancient DNA work is restricted to mtDNA rather than autosomal DNA it is not possible to simply examine the DNA sequence of pigmentation candidate genes from early AMH. Therefore it is necessary to approach the question of pigmentation using alternative methods. The lighter skin pigmentation of chimpanzees is usually attributed to either the effects of living in a dense jungle habitat (where UVR will be lower than in more open environments) or to their covering of body hair, which might help to protect the skin from UVR damage. Relevant questions to ask when considering the ancestral human skin phenotype are when did hominids first leave jungle environments for the more open savannah and when did they lose their body hair. In a study that attempts to address these issues with respect to skin pigmentation Rogers et al. (2004) used *MC1R* sequence data published by Harding et al.

(2000) to try to determine the time since the loss of human body hair. While nonsynonymous changes in the *MC1R* coding region are not common among African populations (presumably due to selective constraint) Harding identified ten such mutations between the last common ancestor of humans and modern chimpanzees [since Harding et al.'s publication three new nonsynonymous mutations in African populations have been identified (John et al., 2003)]. This leads to a relatively high Ka/Ks ratio of 0.63 that Harding et al. feel may best be explained by a scenario in which selective constraint was weak in the distant past but was much stronger more recently (Harding et al., 2000). Rogers et al. suggest that this pattern of variation is consistent with two scenarios. In the first, selective pressure for a darker skin color was weaker in the past due to habitation in dense forest regions where UVR would be weak, regardless of latitude. However since paleoclimatic evidence indicates a shift towards savannah-like environments in Africa around 2.8 MYA (deMenocal, 1995) they feel that this may be unlikely. The second possibility is that selection on *MC1R* was weak in the past because our ancestors were shielded from UVR by a protective covering of body hair similar to that of non-human primates today. Assuming that the loss of body hair in a high UVR environment would have been accompanied by the elimination of variation surrounding *MC1R* nucleotide positions associated with darker skin, Rogers et al. used the sequence data reported by Harding et al. to estimate that humans lost their body hair 1.2 million years ago (Rogers et al., 2004). This date suggests that UVR-mediated natural selection began influencing skin pigmentation in African hominid species as early as *H. ergaster*. This would mean that some of the *H. erectus* species found outside of Africa before this time may have been characterized by relatively hairy bodies and possibly genetic variants

associated with lighter skin pigmentation (similar to chimpanzees). It is difficult to know when or how these early erectus populations would have lost their hair and how that may have led to variation in skin pigmentation genes in these populations.

If natural selection began to have an effect on reducing variation in pigmentation candidate genes as early as *H. ergaster* then it is probable that the first anatomically modern humans (AMH) to emerge in Africa were characterized by dark skin pigmentation. However, we know today that pigmentation shows extensive variation across Africa (Relethford, 2000). Assuming that the levels of pigmentation variation across AMH populations ~ 100 KYA were similar to those observed today in modern African populations we cannot be certain that the first AMH to leave Africa were very darkly pigmented. However, it seems likely that their pigmentation was darker than what is observed in modern populations living in lower UVR regions.

Khoisan-speaking groups are an example of modern sub-Saharan African population with rather light skin (especially relative to their Bantu neighbors). Some studies have positioned Khoisan speaking groups such as the !Kung close to the root of the mtDNA (Chen et al., 2000) and Y chromosome (Semino et al., 2002) human phylogenies. Knight et al. (2003) believe that the deep genetic divergence of the Khoisan and other click-language speakers like the Sandawe and Hata indicate that click languages like those spoken by these groups may have appeared early in the history of modern humans. This places groups like the Khoisan close to the base of the modern human lineage (Knight et al., 2003), making the lighter skin of Khoisan populations and the high *OCA2* 355*A frequency observed in the CEPH diversity panel Khoisan and Moodley et al. samples (Chapter Four) particularly interesting. While today Khoisan

speaking groups inhabit southern Africa (where UVR is lower than in other regions), their genetic relatedness with some Ethiopian groups, archaeological evidence, and the deep separation of Khoisan from other click-speaking groups like the Sandawe and Hata suggest that the Khoisan may originally have inhabited regions further to the north (including modern day Ethiopia and Sudan) where UVR is higher (Nurse et al., 1985). It is currently unknown if the lighter pigmentation of the Khoisan is the result of selective pressures favoring decreased melanin in the lower UVR environments of southern Africa to maximize Vitamin D₃ synthesis or relaxation functional constraints outside of a high UVR environment. Alternatively their lighter pigmentation might represent the retention of a more ancestral phenotypes and pigmentation alleles. Under the first scenario genetic drift could be responsible for the light phenotype observed among the Khoisan. Tests of purifying selection in *MC1R* in Khoisan and sub-Saharan African populations were significant, suggesting that natural selection does constrain variation in *MC1R* in these populations (John et al., 2003). However, as pigmentation is a polygenic trait it may be that the lighter skin pigmentation of Khoisan groups is due to mutations in other genes—including quite possibly *OCA2*. In any event, the Khoisan represent perhaps an example of how light skin pigmentation could have been in some early AMH populations.

A different study that may be used to infer time since loss of body hair in humans examines the origin of human body lice (Kittler et al., 2003). In this study the mtDNA sequence of difference louse species with unique ecological niches on the human body were compared: *Pediculus humanus corporis* and *P. h. humanus* feed on the body but live in clothing while head lice (*P. h. capitus*) live and feed exclusively on the scalp. The authors date the most recent common ancestor of body lice to $\sim 72,000 \pm 42,000$

years ago. Citing the fact that new ecological niches are often exploited quickly they believe that these dates coincide with the start of frequent use of clothing. The mtDNA evidence also shows higher within-Africa sequence diversity and evidence of a recent population expansion that parallel global expansion of modern humans. Taken together these data led the authors to conclude that clothing is a recent human invention and that it may not have been widely adopted until humans moved into cooler climates.

While clothing may have come into widespread use only after humans began exploiting cold environments it is also possible that clothing did not become common until recently because of a protective covering of body hair. If humans retained body hair until late in human evolution then selective pressure to maintain darkly pigmented skin in high UVR regions such as those found in Africa was weak until fairly recently. This would suggest that the ancestral skin pigmentation phenotype may have been more similar to the lighter coloration observed in chimpanzees and that darker skin in tropical populations evolved both recently and quickly. It also suggests that fewer changes were required to produce a light skin pigmentation phenotype consistent with what is observed among northern European populations than if the ancestral condition were very dark skin.

Given the similarities between West Africans and Island Melanesians at *OCA2* and *ASIP* demonstrated in Chapter Four, as well as the strong signals of selection in European populations at *TYR* and *MATP*, I feel that a scenario in which the skin pigmentation of later African hominids was relatively dark is more likely than one in which it was lighter until only fairly recently. While similarities at *OCA2* and *ASIP* between West Africans and some Island Melanesians may be the result of descent from a common darkly pigmented ancestor, Chapters Three and Four also demonstrated that not

all Island Melanesians are similar to West Africans at either of these loci. This may reflect Austronesian influence among Island Melanesian groups, and suggests that Austronesian speakers arriving in Island Melanesia ~3,200 years ago may have been more lightly pigmented than the *in situ* residents that they met. These Austronesians may have undergone a rapid period of adaptation to the high UVR conditions of Island Melanesia that did not involve loci at the same genes contributing to darker pigmentation in the original arrivals to the region. A more detailed study of pigmentation candidate gene variation in Island Melanesia, as well as among other groups found along southern coastal migration routes out of Africa and in the Austronesian homelands will be necessary to explain this situation further. As with the genotype-phenotype association studies, we will need better data on the levels of stratification across Island Melanesia in order to interpret the pigmentary gene patterns.

The strong signals of divergence at *MATP* and *TYR* suggest that European populations experienced strong selective pressure on genes associated with lighter skin pigmentation. While such pressure is consistent with predictions made by the Vitamin D Hypothesis, it is also possible that some other, perhaps unidentified selection pressure (perhaps sexual selection) favoring lighter skin in lower UVR regions would also explain these patterns. Sequencing or constructing microsatellite haplotypes around the *MATP* C374G and *TYR* A192C SNPs might help to date when these alleles rose to such high frequencies in European populations. The pattern of variation observed at these loci in the CEPH-diversity sample set indicates that the light alleles occur at high frequencies only in Europe, and that their frequencies decrease as one moves west and south. Identification of genes that have similar effects on pigmentation and show strong signals

of divergence in East Asian populations would provide further support for the hypotheses that genetic mutations associated with lighter skin in Europeans and Asians have independent origins.

Future Directions

Pigmentation Genotype-phenotype Associations in Island Melanesia

If the goal of identifying genes responsible determining pigmentary variation in Island Melanesia is to be pursued it will be first necessary to develop a panel of markers that is able to control for population stratification in the region. One hypothesized source of this stratification is the admixture that resulted when Austronesian (AN) speakers began arriving in the region 3,200 years ago and met the indigenous, presumably Papuan (P) speaking populations. However, as those indigenous populations had likely been present beginning as early as 40 KYA and subject to inter-population divergence due to drift during that time it is difficult to know what type of admixture processes to model. It is likely that the situation is more complex than a simple two- or three-way admixture scenario. Given the great length of time since these admixture events, and because the phenotypic and genotypic data presented in Chapters Two and Three suggest that geographic differences might be more important than linguistic differences the utility of a panel of markers that can differentiate strictly between AN and P speakers is questionable. It may be more useful to construct multiple panels of markers that can

differentiate between AN and P speakers within a region (or island). This will be more costly and time consuming but may prove to be more effective.

One possible way to rapidly generate these panels would be to type a subset of Island Melanesians on the Affymetrix 100K mapping array chip panel (or some similar system) to generate allele frequency information. Many programs used in admixture mapping to identify population stratification do not require that subpopulations be specified *a priori*. Rather, these programs (such as STRUCTURE or ADMIXMAP) are able to identify subpopulations as well as markers in the dataset that may best discriminate between them without prior information. By identifying alleles that will be helpful in controlling for this structure we will be able to perform genotype-phenotype mapping in the region (Hoggart et al., 2003).

It may also be necessary to identify SNPs in pigmentation candidate genes polymorphic in Island Melanesia by sequencing regions of these genes in a small Island Melanesian sample. This may help to overcome some of the ascertainment bias introduced by using a panel of SNPs that had been primarily discovered in European populations. It may also be more cost-effective to type SNPs identified as polymorphic in the HapMap East Asian and West African populations rather than sequence a large number of genes in Island Melanesians, although SNPs typed in the HapMap also are affected by ascertainment bias.

Identifying Genotype-phenotype Linkage in Other Admixed Populations

Linkage between *MATP* and skin pigmentation was confirmed in a combined sample of admixed African Americans and African Caribbeans. Due to the allelic associations generated during the admixture process, admixed populations can be especially useful in mapping genes associated with traits that differ between the parental populations if population stratification can be controlled for (this process is called admixture mapping). An additional admixed population that may be useful for this work would be the San Luis Valley Hispanics of Colorado (Bonilla et al., 2004). As this population is comprised of individuals with primarily European and Native American ancestry it would provide an opportunity to confirm that *MATP* has an effect on pigmentation in populations that are not the result of European and African admixture alone.

In Chapter Four I demonstrated that the lighter skin of Europeans and East Asians has independent evolutionary origins for at least a subset of the genes determining skin color. *TYR* and *MATP* are both associated with lighter skin in Europeans, but there will be functional variants, perhaps in those genes or perhaps in others, linked with lighter skin in East Asians that may be identified in future studies. Due to the population affinities between Native Americans and East Asians it is possible that admixture mapping studies in Hispanic populations that have primarily European and Native American ancestry may help in identifying linkage between these as of yet unidentified East Asian variants and pigmentation phenotype.

Confirming the Signal of Selection in *MATP* and *TYR*

The strong differentiation between Europeans and all other populations at *MATP* C374G and *TYR* A192C and the confirmation of an admixture linkage between these genes and skin pigmentation suggests that they may have played an integral role in the origins of lighter skin in European populations. In an effort to better understand the timing and origins of lighter skin pigmentation in Europeans it would be extremely useful to be able to date these mutations and chart their expansion. The patterns of allele frequency variation in the CEPH-diversity panel that I demonstrated in Chapter Four suggest that these alleles rose to high frequency as part of a selective event that occurred relatively recently. These alleles occur at high frequency only in Europe, and frequency rapidly declines with increasing geographic distance. While we expect functional constraints on pigmentation to limit the spread of these light alleles to populations living in high UVR regions, the relative absence (or very low frequency) of these alleles in regions of low UVR in East Asia is interesting. This geographic localization implies a recent origin of the light alleles in Europe, as there may not have been time for these alleles to be spread westward via gene flow.

A relatively recent selective event should result in extended LD and an excess of low frequency variants around the functional SNPs. We can measure LD levels by constructing SNP or microsatellite haplotypes. As recent, strong selection should increase levels of LD around the functional SNP we expect that the European variants at these genes will occur on a unique haplotype background relative to other populations. By sequencing these genes we will be able to apply tests of neutrality such as Tajima's D

(Tajima, 1989) and Fu and Li's D^* (Fu and Li, 1993). These tests measure whether the observed frequencies of segregating mutations deviate significantly from what is predicted under the neutral model. Directional selection will lead to an excess of low-frequency variants and will be identified as negative Tajima's D or Fu and Li's D^* values. One limitation to using these test statistics is that population expansion may also produce a similar pattern of sequence variation. Distinguishing between negative values at Tajima's D or Fu and Li's D^* that are due to directional selection at *MATP* and *TYR* from negative values due to a Neolithic European expansion may prove difficult, although it is possible to compare these values to those produced under a simulation that takes expansion into account (Akey et al., 2004). SNP or STR haplotypes may help to resolve the issue, as population expansion should not generate long haplotype blocks. An alternative would be to compare sequence variation at these genes to variation in regions that already take the demographic history of the populations into account (Hammer et al., 2004). Such regions should be located in introns and preferably in high regions of recombination.

Distinguishing Between Sexual and Natural Selection

In Chapter Four I used pairwise locus-specific F_{ST} to identify pigmentation candidate SNPs showing high between-population divergence due to selection. One limitation to this test, and to all sequence-based and LD-based tests of selection, is that it is unable to distinguish between the effects of natural and sexual selection. As both have been put forward as possible explanations for global variation in human skin

pigmentation it would be very interesting to find some way to discriminate between the two types of hypotheses. Some of these limitations in determining the role of sexual selection in shaping broad patterns of pigmentation are discussed below.

The general trend towards lighter skin pigmentation in human females has been cited by some as evidence of sexual selection influencing human pigmentation variation (Van den Berghe and Frost, 1986; Frost, 1988; Frost, 1994). However, proving this is somewhat more difficult. Van den Berghe and Frost (1986) compiled information suggesting a general male preference for lighter females across different human populations. They believe that colonial influences may be a proximate historical explanation for these preferences in many societies, but that preference for lighter females pre-dates European colonial expansion. However, I feel that modern preferences in skin pigmentation should be interpreted with caution when trying to infer how sexual selection may have acted in early human populations to shape global patterns of variation in human skin pigmentation.

Conventionally sexual selection is viewed as being driven by female choice. It is the males of non-human species such as Jackson widow birds and birds of paradise that typically evolve elaborate secondary sexual characteristics, not females. This is in part because females tend to be “choosy” in selecting their partners while males may be less interested in quality. This makes the hypothesis that pigmentation variation in the human species has been driven by males selecting lightly pigmented females unusual to say the least. However, humans do deviate from this general pattern in that males may also be “choosy” when selecting a mate focusing on physical characteristics such as waist-to-hip ratios (Singh, 1993). Aoki (2002) suggested a mechanism explaining how

that male-choice in humans may affect overall female pigmentation in a population based on unequal distribution of wealth. In these societies males with greater access to resources should mate before males with few resources—giving them a wider choice among available females. If these males were to choose based on a physical trait, such as skin pigmentation, it could lead to changes in the phenotype over time (Aoki, 2002).

While Aoki's explanation may be the best put forward thus far to explain how male choice could cause trends associated with lighter pigmentation in females across many populations it does not explain the general correlations of skin pigmentation with UVR. Further, it does not provide insight as to whether or not high pairwise F_{ST} signals have been generated due to natural or sexual selection. Currently I am not sure how to distinguish between the two scenarios using population divergence information based on F_{ST} signals, but I feel that the geographic clines in pigmentation are most consistent with natural selection-based hypotheses. However, sexual selection likely plays an important role in shaping variation in pigmentation within regions.

Conclusions

Skin pigmentation is a trait that shows broad patterns of variation across the human species that have been shaped by natural selection. However, within localized regions like Island Melanesia patterns of variation may be best explained by the action of drift and sexual selection, and not strictly as an adaptation to UVR conditions. Before searching for genotype-phenotype associations in localized populations one should first test for population substructure. If substructure is present then a method controlling for

this stratification while testing for genotype-phenotype associations must be implemented. Care should also be taken when selecting pigmentation candidate SNPs for these studies so as to minimize the effects of ascertainment bias that can be introduced when using SNPs identified in either a small sample or from a single population. Measures of population divergence (e.g. F_{ST}) at pigmentation SNPs provide evidence that selection may have shaped broad patterns of pigmentation variation across the human species. These high pairwise F_{ST} values and the distribution of the light and dark alleles at these loci across the CEPH-diversity sample set provide evidence that darker populations like West Africans and Island Melanesians may be dark in part due to similar mutations at *OCA2* and possibly *ASIP*, while the lighter pigmentation of Europeans may be in part the result of recent mutations in *TYR* and *MATP* that are not shared with East Asians. Future research, perhaps sequence-based, will provide a means to confirm these signals of selection. This work and the future studies of skin pigmentation that it generates will help to demonstrate that skin color variation, rather than being evidence of deep divisions between human populations is instead the result of the shared ability of humans with all living organisms to evolve and adapt to novel environmental conditions.

Literature Cited

- Akey J, Eberle M, Rieder M, Carlson C, Shriver M, Nickerson D, and Kruglyak L (2004) Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biology* 2:1591-1599.
- Aoki K (2002) Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited. *Annals of Human Biology* 29:589-608.
- Bonilla C, Boxill L, Donald S, Williams T, Sylvester N, Parra E, Dios S, Norton H, Shriver M, and Kittles R (2005) The 8818G allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Human Genetics* 116:402-406.
- Bonilla C, Parra E, Pfaff C, Dios S, Marshall J, Hamman R, Ferrell R, Hoggart C, McKeigue P, and Shriver M (2004) Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Annals of Human Genetics* 68 (Pt 2):139-153.
- Chen Y, Olckers A, Schurr T, Kogelnik A, Huoponen K, and Wallace D (2000) mtDNA variation in the South African Kung and Khwe-and their genetic relationships to other African populations. *The American Journal of Human Genetics* 66.
- deMenocal P (1995) Plio-Pleistocene African Climate. *Science* 270:53-59.
- Frost P (1988) Human skin color: A possible relationship between its sexual dimorphism and its social perception. *Perspectives in Biology and Medicine* 32:38-58.

- Frost P (1994) Geographic distribution of human skin colour: a selective compromise between natural selection and sexual selection? *Human Evolution* 9:141-153.
- Fu Y, and Li W-H (1993) Statistical tests of neutrality of mutations. *Genetics* 133:693-709.
- Hammer M, Garriga D, Wood E, Wilder J, Mobasher Z, Bigham A, Krenz J, and Nachman M (2004) Heterogeneous patterns of variation among multiple human x-linked loci: the possible role of diversity-reducing selection in non-Africans. *Genetics* 167:1841-1853.
- Harding R, Healy E, Ray A, Ellis N, Flanagan N, Todd C, Dixon C, Sajantila A, Jackson I, Birch-Machin M, and Rees J (2000) Evidence for variable selective pressures at MC1R. *American Journal of Human Genetics* 66:1351-1361.
- Hoggart C, Parra E, Shriver M, Bonilla C, Kittles R, Clayton D, and McKeigue P (2003) Control of confounding of genetic associations in stratified populations. *American Journal of Human Genetics* 72:1492-1504.
- John P, Makova K, Li W, Jenkins T, and Ramsay M (2003) DNA polymorphism and selection at the melanocortin-1 receptor gene in normally pigmented southern African individuals. *Annals of the New York Academy of Sciences* 994:299-306.
- Kittler R, Kayser M, and Stoneking M (2003) Molecular evolution of *Pediculus humanus* and the origin of clothing. *Current Biology* 13:1414-1417.
- Knight A, Underhill P, Mortensen H, Zhivotovsky L, Lin A, Henn B, Louis D, Ruhlen M, and Mountain J (2003) African Y chromosome and mtDNA divergence provides insight into the history of click languages. *Cell Biology* 13:464-473.

- Mundy N, and Kelly J (2003) Evolution of a pigmentation gene, the melanocortin-1 receptor, in primates. *American Journal of Physical Anthropology* 121:67-80.
- Nurse G, Weiner J, and Jenkins T (1985) The San yesterday and today. In G Harrison (ed.): *The peoples of southern Africa and their affinities*. Oxford: Clarendon Press, pp. 103-126.
- Paabo S, and Wilson A (1991) Miocene DNA sequences--a dream come true? *Current Biology* 1:45-46.
- Parra E, Kittles R, Argyropoulos G, Pfaff C, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey W, Jin L, McKeigue P, Kamboh M, Ferrell R, Pollitzer W, and Shriver M (2001) Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *American Journal of Physical Anthropology* 114:18-29.
- Parra E, Marcini A, Akey J, Martinson J, Batzer M, Cooper R, Forrester T, Allison D, Deka R, Ferrell R, and Shriver M (1998) Estimating African American admixture proportions by use of population-specific alleles. *The American Journal of Human Genetics* 63:1839-1851.
- Post P, Szabo G, and Keeling M (1975) A quantitative and morphological study of the pigmentation system of the chimpanzee with light and electron microscope. *American Journal of Physical Anthropology* 43:435-444.
- Relethford J (2000) Human skin color diversity is highest in sub-Saharan African populations. *Human Biology* 75:773-780.
- Rogers A, Iltis D, and Wooding S (2004) Genetic variation at the MC1R locus and time since loss of human body hair. *Current Anthropology* 45:105-108.

- Rowe N (1996) *The pictorial guide to the living primates*. Charlestown, Rhode Island: Pogonias Press.
- Semino O, Santachiara-Benerecetti A, Falaschi F, Cavalli-Sforza L, and Underhill P (2002) Ethiopians and Khoisan share the deepest clades of the human Y-chromosome phylogeny. *American Journal of Human Genetics* 70:265-268.
- Shriver M, Parra E, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R (2003) Skin pigmentation, biogeographical ancestry and admixture mapping. *Human Genetics* 112:387-399.
- Singh D (1993) Adaptive significance on female physical attractiveness: role of waist to hip ratio. *Journal of Personality and Social Psychology* 65:293-307.
- Sumner P, and Mollon J (2003) Colors of primate pelage and skin: objective assessment of conspicuousness. *Journal of Primatology* 59:67-91.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Van den Berghe P, and Frost P (1986) Skin color preference, sexual dimorphism and sexual selection: a case of gene culture co-evolution? *Ethnic and Racial Studies* 9:87-113.
- White T, Asfaw B, DeGusta D, Gilbert H, Richards G, Suwa G, and Howell F (2003) Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423:742-747.

VITA

Heather Lynne Norton

EDUCATION:

2002-2005 Ph.D. Anthropology. The Pennsylvania State University
2000-2002 M.A. Anthropology. The Pennsylvania State University
1996-2000 B.A. Anthropology. The Pennsylvania State University

AWARDS AND MEMBERSHIPS:

2003-present Wenner Gren Foundation Dissertation Improvement Grant #7138
(\$24,828.00)
2003-2004 Weiss Fellow
2000-2001 Weiss Fellow
Pan-American Society for Pigment Cell Research, member
American Society of Human Genetics, member
American Association for the Advancement of Science, member

SELECTED PAPERS, POSTERS, AND PUBLICATIONS:

Norton HL, Shriver MD, Friedlaender J. (in preparation) Pigmentation and associated candidate gene variation in Island Melanesia. In J Friedlaender (ed.): Population Genetics, Linguistics, and Culture History in the Southwest Pacific: A Synthesis. Cambridge: Oxford University Press

Norton HL, Friedlaender J, Merriwether DA, Koki G, Mgone S, Shriver M. in press: *American Journal of Physical Anthropology*. Skin and Hair Pigmentation Variation in Island Melanesia.

Makova K, and Norton HL. in press: *Peptides*. Worldwide Polymorphism at the *MC1R* locus and normal pigmentation variation in humans.

Shriver M, Parra E, Dios S, Bonilla C, Norton HL, Jovel C, Pfaff C, Jones C, Massac A, Cermaon N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart C, McKeigue P, and Kittles R. 2003. Skin pigmentation, biogeographically ancestry, and admixture mapping. *Human Genetics* 112:387-399.