



The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title: Genomewide Association of Quantitative
Pigmentary Traits in Admixed US
Populations

Author(s): Heather L. Norton

Document Number: 252861

Date Received: April 2019

Award Number: 2013-DN-BX-K011

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Report Title: Genomewide Association of quantitative pigmentary traits in admixed US populations: Final Report
Award Number: 2013-DN-BX-K011
Author(s): Heather L. Norton

Abstract

Pigmentation of the skin, hair, and eyes shows remarkable levels of variation across human populations. Identifying the genetic loci that underlie this variation has potential applications in the field of forensic science, as this may make it possible to determine the probability of certain phenotypic characteristics (e.g. skin, hair, and eye color) using an unidentified sample of DNA. Such information could potentially be useful as a way to direct police investigations by limiting suspect pools in the absence of other intelligence. These predictions may also help to facilitate the identification of missing persons from skeletonized human remains and remains from mass disaster sites. To date most of the research in this field has focused on identifying genetic loci responsible for variation in hair and iris pigmentation in populations of European ancestry. However, far less is known about phenotypic variation and associated genetic loci in non-European and particularly in admixed populations. A thorough understanding of both phenotypic and genetic variation in these populations will be essential in order to determine if prediction of pigmentary phenotype from DNA is possible, particularly across populations of diverse origin.

The goals of this project were to address this problem by characterizing phenotypic and genetic variation in a diverse set of populations and identifying associations between genotypes and quantitatively measured pigmentary phenotype. Pigmentation measurements and DNA were collected from individuals from Cincinnati, OH who self-identified as African American or Hispanic. We supplemented our sample to include two previously collected admixed samples from outside of the United States (Mexico and Brazil) as well as a previously collected sample of individuals who self-reported as being of European, East Asian, or South Asian ancestry from Toronto, Ontario, Canada. Skin, hair, and iris phenotype were assessed using quantitative methods (reflectance spectroscopy and high resolution iris imaging). Dense genotype data was obtained via either the Illumina MEGA or Affymetrix Axiom Array. After phasing, untyped markers are imputed using the 1000 Genomes Panel populations as a reference. Population stratification, which if uncorrected can lead to spurious associations, is assessed using EIGENSOFT. Tests for associations between quantitative pigmentary phenotype and genotype are carried out, using sex and significant ancestry components as covariates.

Our results show that although significant differences in skin pigmentation phenotype exist between our population samples, there is also extensive variation within these samples, particularly in admixed populations. While hair pigmentation is less variable outside of European populations, we demonstrate within-population heterogeneity in our Hispanic sample and variation between Hispanics and other groups. Finally, we observe significant differences in quantitatively assessed iris pigmentation (measured in CIELab color space) among these populations, as well as substantial variation within population samples, including within admixed groups. This highlights the ways that broad categorical phenotyping terms (e.g. “brown”) can mask phenotypically relevant variation.

In a small genotype-association study using European samples we confirm previously reported associations between *SLC45A2*, *IRF4*, and *HERC2* and skin pigmentation and between *IRF4*, *SLC24A4*, *SLC45A2*, and *OCA2/HERC2* and hair

pigmentation. In a GWAS of East Asian skin and iris pigmentation we confirm the importance of the *OCA2* locus in shaping pigmentation phenotype in East Asian populations. We also demonstrate that markers traditionally associated with pigmentation variation in European populations may be unhelpful in predicting pigmentation variation in this group. Looking ahead to the completion of the genetic analyses of our admixed samples, we predict that newly discovered markers in populations of African descent may play a significant role in shaping the genetic architecture of pigmentation in some of these groups. This suggests that the prediction of pigmentary traits from genetic data, particularly in admixed or non-European populations may present a significant challenge.

Table of Contents

Title Page	i
Abstract	ii
Table of Contents	iv
Executive Summary	1
Introduction.....	6
Methods.....	10
Results.....	14
Conclusions	25
References	30
Dissemination of Research Findings	34

Executive Summary

Problem. Pigmentation of the skin, hair, and eyes are easily visible phenotypic traits that vary extensively within and between human populations. As a complex trait, pigmentation is controlled by multiple genetic loci of varying effect sizes. The ability to use genetic data to predict the pigmentary phenotype of an unknown individual has potential applications in the field of forensic science. While this information cannot be used in the same manner that a traditional 13-marker STR profile may be used (e.g. to match a suspect to a crime scene sample or to establish/exclude paternity), it may still provide useful information that can be used to guide an investigation, particularly in early stages or when other evidence is limited. For example, such information may help to direct early stages of criminal investigations in the absence of other intelligence (such as reliable eye-witness accounts), or may be used to help facilitate the identification of unknown individuals from skeletonized remains or mass disaster sites.

The foundation for our current understanding of pigmentation genotype-phenotype relationships has largely been based on studies of populations of European ancestry. Specifically, past work has focused on genes that regulate skin, hair, or iris variation within European populations or that can explain mean differences between European and non-European individuals. This strong body of work has led to the identification of loci such as *SLC24A5*, which can explain a large portion of variation observed between populations of European and African ancestry, and the *HERC2/OCA2* complex, which plays an important role in determining blue vs. brown eye color phenotype in European populations. Expanding on such discoveries, efforts have been made to develop panels of genetic markers that can be used to predict skin, hair, and iris pigmentation.

While some of these panels have been forensically validated, there remains some question about their utility when applied in non-European or admixed populations. Concerns with the portability of such tests to other populations groups rests on two primary issues. The first is that many of these studies utilize a categorical system to describe variation in pigmentation phenotype. While such systems have the advantage of being relatively simple and easy to assess, they often mask the true extent of variation in pigmentation phenotype, which is better described as a quantitative, rather than categorical, trait. The second disadvantage is that these panels have been developed largely (but not exclusively) using loci known to play significant roles in regulating European pigmentation variation. This raises the question of whether or not such loci will be relevant in broader and more diverse population samples, and in particular in populations of admixed origin—such broad applicability is an important component of any marker panel to be used in the US. For example recent research exploring the genetic architecture of pigmentation in African populations suggests that alleles not common in European populations may play an important role in regulating pigmentation in populations there and in other regions. A second research report, focusing on the genetics of pigmentation in Latin American populations identified several novel loci associated with skin and iris pigmentation, including a locus (MFSD12) associated with lighter skin color that is not associated with European ancestry. Such loci would certainly impact pigmentation prediction in admixed populations common in the US. Finally, there is also evidence that the effects of certain

loci on pigmentation phenotype may vary between populations, suggesting that as-yet unknown loci may modulate the effects of key pigmentary predictor SNPs in different genomic backgrounds. Taken together these reports suggest that more thorough investigations using more diverse populations and quantitative phenotyping methods are going to be necessary to identify panels of markers that may be used to predict pigmentation phenotype in diverse population settings.

Purpose. This study set out to address these issues by pursuing two primary goals. The first was to use quantitative methods to assess variation in skin, hair, and iris pigmentation in populations of diverse origin, particularly in admixed populations (individuals who self-identify as African American or Hispanic in the US). This is crucial for understanding not just patterns of variation between, but also within, populations. Such variation can be masked when individuals are placed into broad pigmentation phenotype categories (e.g “blue”, “brown”, “green” for eye color), making it more difficult to identify loci that have small to moderate effects on phenotype. The second goal was to identify genetic loci associated with this quantitatively assessed variation using dense genotype arrays that to include variants of moderate to low frequency in non-European populations.

Research Design. The original goal of this study was to focus on a large sample of admixed individuals (500 self-identified African Americans and 500 self-identified Hispanics) from Cincinnati, OH. Due to lower-than expected participation (259 African Americans, 90 Hispanics) we augmented these admixed samples with 101 individuals from Palenque Mexico and ~480 Brazilians from Sao Paulo. We also incorporated previously collected data from 583 European, 425 East Asian, and 314 South Asian individuals from Toronto, Ontario, Canada. In addition to obtaining quantitative phenotypic measurements and a saliva sample to be used for DNA analyses, participants also completed a questionnaire providing basic information about their ancestry and self-reported descriptions of their skin, hair, and iris pigmentation.

Constitutive skin pigmentation was measured in the inner upper arm using a DSM II ColorMeter (Cortex, Technology, Hadsund, Denmark) following established protocols. Pigmentation values were recorded as Melanin Index (MI) and in CIELAB coordinates. Higher MI values correspond to darker skin colors. The CIELAB color system describes color using three dimensions: L^* (measuring brightness), a^* (measuring the intensity of green/red reflected light), and b^* (measuring the intensity of blue/yellow reflected light). The DSM II was also used to record hair pigmentation, with measurements taken at the crown of the head. Three measurements were taken at each site with mean skin and hair values calculated for each individual. Iris pigmentation was assessed using high-resolution digital images of the iris using a Miles Research Professional Iris camera under controlled lighting conditions. Iris images were analyzed using a custom computer program developed in the lab of the co-PI (<http://iris.davidcha.ca>). Briefly, the pupil and sclera were cropped from the image and a wedge was extracted from the iris. Average L^* , a^* , and b^* values were calculated for pixels within the wedge to provide quantitative measurements of iris pigmentation.

2mL of saliva were collected from participants and DNA was extracted following standard protocols. DNA samples from the African American, Hispanic, European, East

Asian, and South Asian samples were genotyped on the Illumina Infinium Multi-Ethnic Array, which includes 1.7 million markers targeting common variants across diverse population groups. Genotyping of the Brazilian samples using this array is currently in progress. The Mexican samples were genotyped on the Axiom BioBank array. All genotype data went through standard QC procedures. Untyped markers are imputed using the 1000 Genomes populations as reference panels. Because population structure can lead to false-positive associations, population structure is assessed using the program EIGENSOFT, and significant ancestry components are used as covariates in association analyses.

Findings. Our quantitative measurements of skin pigmentation reveal significant variation within and between populations in MI. As expected, Europeans exhibit the lowest and least variable MI values (mean = 35.3, range: 28.6-58.5), with African Americans (mean = 63.1, range: 33.4-104.7) and South Asians (mean = 47.5, range: 32.0-68.7) exhibit the highest and most variable values. Mean skin MI was significantly different among the populations ($F = 1044$, $p < 2 \times 10^{-16}$), however the MI distributions of all populations overlapped. In other words, certain ranges of MI are common to all populations examined here. A pairwise t-test (with Bonferroni correction) comparing skin MI among pairs of populations revealed that all pairs of populations differed significantly from each other, with the exception of the Hispanic-East Asian (mean Hispanic MI = 39.2, mean East Asian MI = 37.9) and Mexican-South Asian (mean Mexican MI = 49.0, mean South Asian MI = 47.5) comparisons. As expected, admixed populations exhibited a wide range of MI values. MANOVA comparing the skin CIELab values confirmed that all populations differed from each other in CIELab space as well ($p < 0.0001$).

In contrast to skin pigmentation, Europeans had the lowest hair MI values and highest CIELab values, while also exhibiting the widest range of variation. This confirms visual assessments and the wide range of self-reported hair colors recorded by European participants in their participant questionnaires. Populations were significantly different from each other in both MI ($p < 2 \times 10^{-16}$) and CIELab ($p < 2 \times 10^{-16}$) coordinates, although as with skin color all populations overlapped in hair MI and CIELab values. Pairwise comparisons showed that the European sample was significantly different from all other populations, as was the Hispanic sample (which had the second lowest mean hair MI value: 149.8).

Iris pigmentation (as measured in CIELab space) was compared among populations. The admixed African American and Hispanic population samples tended to exhibit darker eye colors compared to European individuals, with lower mean L^* , a^* , and b^* values. However, these values were highly variable in both admixed samples. While the African American and Hispanic irises primarily occupy color space associated with dark and light brown irises in the European, East Asian, and South Asian samples, they also included individuals with lighter and intermediate eye colors, including one blue-eyed individual.

As genetic analyses in the admixed populations are still currently underway no conclusions can be drawn about specific associations between particular loci and quantitative pigmentation phenotypes. However, we have completed a preliminary investigation to identify associations between nine candidate SNPs and skin and hair pigmentation in the European sample as well as a full genome-wide association study to identify loci associated with skin and iris variation in the East Asian sample.

We confirmed the association of three alleles (rs16891982 in *SLC45A2*, rs12203592 in *IRF4*, and rs12913832 in *HERC2*) with lower MI values in the European sample. We also confirmed the association of four alleles (rs12203592 in *IRF4*, rs12896399 in *SLC24A4*, rs7495174 in *OCA2*, and rs12913832 in *HERC2*) with variation in hair L^* , a^* , and b^* values. Interestingly, we determined that the derived T allele at rs12203592 is associated with lighter skin color but darker hair color in this population, indicating that this SNP has different effects on pigmentation phenotype depending on the tissue in which it is expressed.

In a GWAS of skin and iris pigmentation in the East Asian sample no associations reaching genome-wide significance were identified for skin pigmentation (measured as MI), despite having reasonable power to do so. However, one suggestive signal at rs2373391 (located in the *ZNF804B* locus on chromosome 7) was confirmed in a replication sample, identifying a new locus to pursue. A single SNP in the *OCA2* gene, rs76930569 was significantly associated with two dimensions of eye color (a^* , and b^*). The association is likely driven by LD between this SNP and the nonsynonymous *OCA2* polymorphism rs1800414, previously associated with skin pigmentation variation in East Asian populations.

Conclusions. Although genomic analyses of the admixed samples is still underway, multiple conclusions can be drawn from the work completed to date. First, our use of quantitative, rather than categorical, measurements of skin, hair, and iris pigmentation was crucial for the accurate characterization of both between and within population variation. While we observed significant differences in both skin and hair pigmentation between study populations, we also noted that distributions of MI and CIELab values overlapped for all groups, including those with the most discordant values (e.g. skin MI in African Americans and Europeans). As we expected, we observed significant heterogeneity in skin, hair, and iris pigmentation in our samples, and particularly within the admixed and South Asian population samples. This variation likely highlights both underlying population substructure and differences in admixture proportions as well as potentially novel combinations of pigmentation alleles. For example, the South Asian sample here includes individuals who identify as being from Bangladesh, Pakistan, India (itself highly structured), and Sri Lanka, and the Hispanic sample includes individuals claiming ancestry from Mexico, the Caribbean, and other parts of Central and South America. This highlights the importance of being able to identify and control for population substructure in subsequent GWAS analyses, as failure to do so may result in spurious associations that reflect differences in ancestry proportions rather than true associations.

Our use of quantitative methods to characterize pigmentation traits also supports previous arguments that pigmentation should be considered as a continuous, rather than categorical, phenotype, both to accurately report within population variation as well as to increase power to identify loci of small to moderate effect. This is particularly true for eye color, as we were able to distinguish between different shades of brown eyes (dark vs. light) in CIELab space. This ultimately became important in the identification of loci that have subtle, but important effects on iris pigmentation, as seen in the GWAS of skin and iris pigmentation in the East Asian sample.

Finally, despite having good power (> 0.8) to identify loci explaining 12.5% of variation in skin pigmentation phenotype in the East Asian sample, we identified no candidates reaching genome-wide significance. This may be due to a number of factors, including small sample size, the possibility that loci with important roles in East Asian pigmentation have become fixed (much like the derived allele at rs1426654 in Europeans), or because pigmentation variation in East Asian populations is regulated by many loci of small effect, which may not be common in European populations. Recent studies in African populations have identified alleles impacting pigmentation phenotype that are uncommon in European populations. Taken together, these results suggest that many of the loci currently identified as good predictors of pigmentary traits may underperform in pigment prediction in non-European populations, suggesting that more in-depth investigations into such populations will be key to developing a better understanding of the genetic architecture of pigmentation outside of Europe.

Implications for Policy and Practice. In this study we have prioritized the use of quantitative measurements of pigmentation traits over simpler categorical descriptions. We believe, as others have shown, that treating pigmentation as a quantitative trait improves the ability to identify genetic loci with moderate effects on phenotype, and that such quantitative methods should be used in GWA studies when possible. We also stress that treatment of pigmentation phenotypes as continuous, rather than discrete, traits may help to counter perceptions that individuals belonging to a particular socially defined racial or ethnic group can be distinguished on the basis of pigmentary phenotype. While we observed significant differences among the groups in this study, we also observe notable overlap in the distribution of skin, hair, and eye color values, indicating these groups can be placed into discrete pigmentation categories that neatly correspond to social/ethnic identities (which may themselves be fluid in different contexts). The range of variation that we observed in skin, hair, and eye color within each of these groups also argues against ideas that such groups are homogeneous with respect to pigmentary traits. This fact should be emphasized any time that FDP methods are employed, to prevent investigators from making possibly erroneous links between pigmentation phenotype and race or ethnicity.

While we advocate for the treatment of pigmentation as a continuous trait, we note that translating this information in practical ways to investigators more familiar with categorical treatments of pigmentation (especially for skin and iris color) presents substantial challenges. As a first step in understanding how such predictions would be used by investigators, it will be important to develop a better understanding of how numerical estimates of pigmentation variation correlate with commonly used categorical terms for relevant phenotypes. Training investigators into the highly variable nature of pigmentation phenotypes is recommended.

Finally, while some efforts have been made to demonstrate that current pigmentation prediction panels are applicable to a broad range of populations, we believe that this warrants further investigation. Our work and that of others suggests that the genetic architecture underlying pigmentation variation in European populations (relatively few loci of large effect) may not translate well to non-European populations, where many loci (some that may not be common in Europeans) with smaller effects may contribute to within population variation. Completion of the genomic analyses of our

admixed samples will certainly contribute to this body of knowledge, and we are eager to report those results. For now we recommend that existing pigmentation prediction panels be used with caution in US settings until it can be better determined if loci included in those panels influence pigmentation phenotype in the same way across a broader sample of populations.

Introduction

Over the last two decades there have been several important advances in our understanding of the genetic basis of non-pathological variation in human skin, hair, and iris pigmentation. The identification of loci that impact these externally visible characteristics (EVCs) presents new opportunities for the development and improvement of forensic DNA phenotyping (FDP) techniques. These techniques could be useful in guiding police investigations by limiting suspect pools when additional intelligence is limited (Kayser and de Knijff 2011; Kayser and Schneider 2009; Keating et al. 2013; Walsh et al. 2012) and may also help in the identifying of missing persons and disaster victims. Until recently the bulk of research in this area focused on populations of European ancestry. While this led to the identification of loci that may play significant roles in variation of hair and iris phenotype within such populations, large gaps in our understanding of pigmentation in non-European and admixed populations remained. These gaps limit the application of FDP techniques to the general US public, which includes individuals with ancestry from many different parts of the world or who may self-identify as belonging to an admixed population (e.g. African Americans or Hispanics). Understanding the extent of variation in pigmentary phenotypes between and *within* different populations and the genetic loci that underlie that variation are crucial steps in understanding if it possible to develop a forensic panel that could be used to predict pigmentary phenotypes across an increasingly diverse US population. This project sets out to address these gaps by quantitatively assessing skin, hair, and iris pigmentation in multiple populations, including US individuals who self-identify as belonging to one of two admixed populations (African American or Hispanic). Using dense genome-wide genotype data our second goal is to identify genetic loci associated with the observed phenotypic variation. A subset of loci identified here may ultimately be used to develop panels of markers that can develop predictions of pigmentary phenotype.

Skin, hair, and iris pigmentation are primarily determined by the quantity, type, and distribution of the biopolymer melanin. Melanin production takes place in melanocyte cells, found in the basal layer of the epidermis, hair bulb, and iris. It is produced in cellular organelles known as melanosomes, which are then transported to keratinocyte cells. In the presence of cysteine, alkali-soluble red-yellow pheomelanin is produced, while the absence of cysteine results in the production of brown-black eumelanin. Pigmentation phenotype is determined by the ratio of pheomelanin to eumelanin, total melanin content, and the distribution of melanosomes.

Pigmentation phenotype has been characterized a number of ways. In many studies skin pigmentation has been described in terms of Fitzpatrick Skin Type Classification System (Fitzpatrick 1988), and hair and iris color have been commonly described using descriptive terms (e.g. “blue”, “green”, “brown”) (Frudakis et al. 2007; Frudakis et al. 2003; Han et al. 2008; Sulem et al. 2008; Sulem et al. 2007). Such characterization methods have the advantages of being relatively simple (no special

equipment is required to collect these data) and because they can be collected relatively easily (often through self-report). However, a disadvantage of these simple descriptive phenotypic categories is that they may mask the continuous nature of pigmentary traits (both within and between different populations). In addition, the Fitzpatrick Skin Type Classification System, while relatively simple to use, was originally derived to assess skin melanoma risk in light-skinned populations (Fitzpatrick 1988), and not as a tool to accurately describe skin pigmentation variation. While it has been adapted to include categories for darker skin types, some studies suggest that the wording of questions in the assessment may be culturally biased and/or not useful in characterizing pigmentation variation outside of lightly pigmented, (largely European) populations (Eilers et al. 2013; Pichon et al. 2010). These issues point to a need for a quantitative way to assess skin, hair, and iris pigmentation, both to accurately describe the extent of variation in these phenotypes but to also facilitate the discovery of genetic loci that influence this variation.

While reflectometers have been used to quantitatively assess skin pigmentation since the 1960's and 1970's (Parra 2007), it was only with the more recent development of narrow-band reflectometers that it became possible to directly assess chromophores of primary relevance to skin pigmentation phenotype (i.e., melanin). For example, narrow-band reflectometers, such as the Deraspectrometer (Cortex Technologies, Hadsund, Denmark), utilize the differential absorbance properties of the two primary pigments of the skin—hemoglobin and melanin—at long (red) and short (green) wavelengths (Diffey et al. 1984). By calculating differences of these absorbance spectra the Deraspectrometer yields an estimate of melanin content known as the M Index (MI). One advantage to using MI as a metric is that it is easy to interpret: higher MI values correspond to darker pigmentation. Other quantitative assessments of color have also been applied to skin pigmentation studies, most notably the CIELAB system, which summarizes color differences across three dimensions pertaining to brightness (L^*), and the intensities of red/green (a^*) and blue/yellow (b^*) light. While this system is less intuitive than MI, it provides a more nuanced understanding of color appearance, and may be more suitable for assessments of hair and iris pigmentation. Assessments of skin pigmentation using both types of systems have been used in studies of pigmentation (in both European and non-European populations) for the past fifteen years (Akey et al. 2001; Andrade et al. 2017; Beleza et al. 2013; Bonilla et al. 2005; Bonilla et al. 2004a; Bonilla et al. 2004b; Crawford et al. 2017; Edwards et al. 2010; Hernandez-Pacheco et al. 2017; Jonnalagadda et al. 2016; Martin et al. 2017; Norton et al. 2016; Norton et al. 2006; Shriver and Parra 2000; Shriver et al. 2003; Stokowski et al. 2007) and more recently instruments have been developed that can take measurements in multiple color systems (e.g. the DSMII).

While reflectometers have recently been used to assess hair pigmentation variation, the majority of studies have relied on broad descriptive terms (e.g. “red”, “brown”, “blond”, etc.) to classify hair color, either through self-assessment or via rating by a trained grader. With a few exceptions (Norton et al. 2016; Norton et al. 2006) there is also a pronounced bias towards the study hair pigmentation variation in populations of European ancestry, where there is the broadest distribution of phenotypic variation. Similarly, iris pigmentation has also typically been described in broad categorical terms, often self-reported by study participants. Significant improvements in the objective and

quantitative assessment of iris pigmentation accompanied the development of high-resolution photographic tools to capture iris images (Edwards et al. 2016; Edwards et al. 2011; Liu et al. 2010). Studies using these methods reported extensive variation in iris pigmentation that transcended the traditional categories of “blue”, “green”, “brown”, as well as identified subtle shifts in iris pigmentation across CIELab space in populations commonly thought to be relatively homogeneous in iris pigmentation (Beleza et al. 2013; Edwards et al. 2016; Edwards et al. 2011). Finally, Liu et al. (Edwards et al. 2016; Edwards et al. 2011; Liu et al. 2010) demonstrated that genome-wide association studies of iris pigmentation had improved power when using quantitative, rather than categorical, iris characterizations.

Traditional STR profiling methods are most useful when evidentiary DNA can be compared to a sample from a known individual or profile in a forensic database. While such methods can be used to exclude a suspect from investigation, the use of DNA to predict EVCs has a different application. Genetic predictions of EVCs may help to support or refute eyewitness accounts of a suspect. This can have a potentially significant impact on the early stages of an investigation, as such accounts can be highly varied and sometimes erroneous (Spinney 2008; Wells et al. 2000). Genetic prediction of EVCs may also be useful in reconstructing the appearance of a deceased individual for whom many physical characteristics cannot be determined—this could include skeletonized individuals or fragmentary remains (such as those found at mass disaster sites). While this information cannot be used to make a conclusive identification, they may be helpful in narrowing the pool of potential candidate individuals.

To that end forensic panels have been developed to predict iris, hair, and skin pigmentation. While early attempts to predict hair color focused largely on the associations of the *MC1R* gene with red hair color (Grimes et al. 2001), more recently attempts have been made to develop tests more appropriate to populations in which red hair is not common. While these panels have been successful in distinguishing between qualitative hair colors in European populations (Branicki et al. 2011; Keating et al. 2013; Walsh et al. 2013), their utility in admixed populations remains unknown, due in part to poor understanding of the extent of variation in hair pigmentation phenotype in these groups. Panels developed to predict iris color have built on early work identifying SNPs in the *HERC2/OCA2* complex as being associated with blue vs. brown eye color phenotype (Eiberg et al. 2008; Frudakis et al. 2007; Frudakis et al. 2003; Kayser et al. 2008; Sturm et al. 2008). SNPs from these studies have been included in forensically validated assays (e.g. IrisPlex and HIrisPlex) with the ability to predict blue eye color with an accuracy of 90% (Walsh et al. 2011a; Walsh et al. 2011b; Walsh et al. 2013; Walsh et al. 2012).

A key question regarding such panels is their applicability in non-European populations. This is particularly challenging given the limited number of studies exploring variation in hair or iris phenotype in non-European and/or admixed populations. Attempts to support the use of these panels globally have sometimes focused on the genotype distribution of panel SNPs across the HGDP-CEPH populations (Walsh et al. 2011b). For example, when six loci from the Irisplex test were genotyped in these populations they rarely predicted blue eye color in populations outside of Europe, the Middle East, and Western Europe. However, because the

HGDP-CEPH populations do not have any associated phenotypic information it is difficult to determine how well such panels may predict intermediate eye colors in global populations, or how sensitive they may (or may not) be to more subtle shifts in iris pigmentation. As a specific example, Edwards et al. (Edwards et al. 2016) identified several South Asian individuals homozygous for derived alleles at the SNP rs12913832 (the polymorphism in the *HERC2/OCA2* complex often associated with blue vs. brown eye color) whose eyes were intermediate (but not blue) in color. These authors suggest that other loci may modify the effects of rs12913832 in European and South Asian populations, leading to its different phenotypic effects in each population. While forensic SNP panels rely on multi-locus genotypes for phenotype prediction, the differing effects of rs12913832 in Europeans and South Asians suggests that some caution may be necessary before broadly applying panels developed with Europeans to other groups. On a more positive note, attempts to correlate genotype with iris phenotype in admixed populations from Brazil have met with some success, demonstrating that the *HERC2/OCA2* complex plays a role in determining broad differences in eye color in at least some admixed populations (Andrade et al. 2017; Freire-Aradas et al. 2014).

Because hair and eye color can potentially be temporarily modified (through the use of bleach, hair dye or colored contacts), the ability to develop a forensic panel that can predict skin pigmentation is also of great interest. In 2010 Valenzuela et al. identified a panel of 5 SNPs that could explain ~76% of variation in total hair melanin, ~45% of variation in skin pigmentation, and ~75% of variation in iris pigmentation in an ethnically diverse sample from Arizona (Valenzuela et al. 2010). While this study was significant for not exclusively focusing on individuals of European ancestry, failure to control potential effects of population structure in the sample may have confounded their results. Subsequently Maroñas et al. used African, European, and admixed individuals to develop a Bayesian classifier model based on a panel of 10 markers that could be used to predict skin phenotype into one of three categories: white, intermediate, or black (Maronas et al. 2014). While this study was notable for specifically focusing on admixed samples, its relatively small sample size and reliance on SNPs with a strong European ascertainment bias are potential limitations.

More recently Walsh et al. set out to develop a panel of markers that could be used to predict skin color, with the explicit goal of being able to predict variation within and between continents (Walsh et al. 2017). To that end they used a large sample of individuals, primarily sampled from across Europe, but also including US individuals of diverse backgrounds. Individuals were placed into 5-tiered (very pale, pale, intermediate, dark, and black) and 3-tiered (light, dark, dark black) skin classification categories. As with previous iris studies, unphenotyped individuals from the HGDP-CEPH populations were also included. The authors developed a multiple-linear regression model using 36 SNPs to predict skin pigmentation according to the 5-tiered and 3-tiered models. Their system outperformed that of Maroñas, and allowed them to distinguish between pigmentation variation within continents (primarily relevant to variation within Europeans) as well as between continents. However, the use of the unphenotyped HGDP-CEPH populations as stand-ins for those with a “dark black” pigmentation phenotype raises issues similar to those discussed above for the validation of the IrisPlex panel. Quantitative studies of pigmentation variation in both African and Melanesian populations indicate that skin pigmentation phenotype can be

quite variable in these regions (Crawford et al. 2017; Martin et al. 2017; Norton et al. 2006), and that not all individuals are good models for the “dark “black” skin category used here. As such, it’s difficult to know how well this panel may predict variation in skin pigmentation, particularly across darker skin tones.

More recent studies of pigmentation variation in admixed and African population suggest that models predicting skin and iris pigmentation may need to be modified, due to the discovery of alleles affecting pigmentation variation in African populations that are rare or absent from European-derived populations. For example, in their study of pigmentation variation in KhoeSan populations of Africa, Martin and colleagues identified both classic (*SLC24A5* and *TYRP1*) and novel (*SMARCA2/VLDLR* and *SNX13*) loci influencing pigmentation in these groups (Martin et al. 2017). Their work suggests that pigmentation variation in these, and likely other African populations, is regulated by a large number (> 50) of loci with relatively minor effects—this contrasts with current models of understanding (based on studies of European populations) that suggest pigmentation is regulated by relatively few genes of major effect. Their findings are bolstered by a larger study of African pigmentary diversity conducted by Crawford et al. (Crawford et al. 2017), who also identify novel loci (*MFSD12*, *DDB1*, *TMEM138*) in addition to others (*SLC24A5*, *OCA2*, *HERC2*) influencing pigmentation phenotype across the continent. In a study of admixed Puerto Rican and African American individuals in the US, Hernandez-Pacheco, et al. identify associations between the loci *BEND7* and *PRFP18* and quantitative estimates of skin color (Hernandez-Pacheco et al. 2017). A recent GWAS study investigating skin and iris pigmentation in a sample of ~ 6,000 Latin Americans identified several novel loci associated with pigmentary traits. These include a mutation in the *MFSD12* gene that is associated with lighter skin color. This mutation is thought to be related to East Asian/Native American (rather than European) ancestry (Adhikari et al. 2019). Taken together, these results suggest that pigmentary architecture, particularly in non-European and admixed populations, may be more complex than previously appreciated. The development of broadly applicable forensic panels that could be used to predict different pigmentary phenotypes should therefore include the use of diverse populations with associated quantitative phenotypic data.

Given the desired goal of developing predictive models of pigmentary phenotypes across a range of populations, particularly those relevant in a US context, this project set out to conduct GWA studies using populations of diverse origins, including two US admixed populations (African Americans and Hispanics). Two advantages to including admixed populations in pigmentation GWAS include the anticipated wide range of phenotypic variation observed in these groups and the ability to utilize admixture mapping approaches to identify relevant genetic loci.

Methods

Participant Recruitment and Consent Process. The original goal of this study was to recruit a total of 1000 admixed individuals from the Cincinnati, OH region: 500 self-identified African American individuals and 500 self-identified Hispanic individuals. It should be noted that while we specifically advertised for participants that self-identified as “Hispanic”, a small number (seven) of participants reported their ancestry as “Hispanic/Latino”. Further review of questionnaire responses indicated that three of

these individuals also included at least one parent/grandparent who was Brazilian (e.g. Latino). We have included these individuals into our broader “US Hispanic” sample. Participants in this broad “US Hispanic” sample report ancestry from a range of countries, including Puerto Rico Mexico, Columbia, Cuba, and Venezuela among others. Participants were recruited through the University of Cincinnati community, and later more broadly through the city of Cincinnati, via recruitment flyers and course announcements. Recruitment materials were also translated into Spanish in an attempt to boost US Hispanic enrollment numbers. Recruitment of both African American and US Hispanic individuals proved to be far more difficult than expected, due we believe in part to external events that occurred in the Cincinnati community (as noted in bi-annual reports). All individuals recruited to the project were provided with an informed consent document, which was reviewed with study personnel. This document, as well as study procedures, were approved by the University of Cincinnati’s IRB (UC IRB #2013-7610). A Change of Scope (submitted in July 2014) specified the inclusion of an additional ~1500 individuals of European, East Asian, and South Asian descent. Phenotypic data from these individuals had already been collected by the co-PI, and DNA samples were available to utilize in genome-wide association testing for pigmentary traits.

In an effort to boost the size of our admixed sample, we reached out to other researchers with similar datasets. This resulted in the inclusion of data from 101 Mexican individuals from the city of Palenque collected by Dr. Abigail Bigham at the University of Michigan. Phenotypic data for these individuals included skin and hair reflectance measurements and DNA genotype data (see below). We also were able to extend our collaboration to include ~480 admixed individuals collected by Dr. Celso Texiera Mendes Junior of the University of Sao Paulo, Brazil. Phenotypic data from these individuals includes skin reflectance, categorical iris assessment, and categorical descriptions of hair color. DNA samples from these Brazilian individuals are also available and are in the process of being genotyped on the Illumina MEGA Array. Due to delays in the approval for the inclusion of these data, and in the final purchase of genotyping kits to be used on these samples genotyping is still in process. Because of the different nature of data collected from each group, as well as the different stages of genomic analyses, summary tables have been provided at the end of this section to denote the types of phenotypic data available for each sample and the stage of genotypic analyses. In subsequent discussion, tables, and figures, the “US Hispanic” sample refers to individuals sampled in Cincinnati, OH who self-identified as being either Hispanic or Latino. “Mexican” refers to the Palenque sample collected by Dr. Bigham, while “Brazilian” refers to the samples collected by Dr. Texiera Mendes. Questionnaire data. US (African American and US Hispanic) and Canadian (European, East Asian, and South Asian) participants completed a brief questionnaire prior to phenotypic data collection. This provided information about their self-reported ancestry, place of birth and first language of parents and grandparents. The questionnaire also asked a series of questions about pigmentation and skin sensitivity to sun, allowing us to place individuals into one of the six Fitzpatrick skin type categories.

Phenotypic measurement. Constitutive skin pigmentation (pigmentation in unexposed areas of the skin) was measured in the inner upper arm using a DSM II ColorMeter (Cortex Technology, Hadsund, Denmark). The DSM II ColorMeter measures skin reflectance by shining a light on the surface of the skin and measuring the amount of

light that is reflected back at different wavelengths, providing a non-invasive and quantitative method to investigate skin color. The probe of the DSM II was placed gently against the surface of the skin for ~ 5-7 seconds to allow the light emitted by the DSM II diode to shine on the skin and for reflectance measurements to be obtained.

Reflectance data was recorded as Melanoma (M) and Erythema (E) index, CIELab values (L^* , a^* , and b^*), and RGB values. Three measurements from the arm were taken, and averaged together to yield mean skin reflectance values. Analyses focus on the M, L^* , a^* , and b^* measurements.

Hair pigmentation was quantitatively assessed in all subjects using the DSM II. The probe of the DSM II was placed against hair at the crown for 5-7 seconds to allow the light emitted by the DSM II diode to shine on the hair and for reflectance measurements to be taken. As with skin, three measurements were taken and averaged together to produce mean MI and L^* , a^* , and b^* values. Hair that was dyed or graying was not measured. We also asked for optional hair samples from participants to allow for quantitative assessments of melanin content by Dr. Kazu Wakamatus at the Fujita University School of Health Sciences. The majority of our participants declined to provide a hair sample.

A photograph of each subject's iris was taken with the Miles Research Professional Iris camera (Miles Research, California, USA). This camera consists of a Fujifilm Finepix S3 Pro 12-megapixel DSLR mounted on a Nikkor 105mm macro lens. A coaxial biometric illuminator delivers a constant and uniform source of light to each iris at 5500K (D55 illuminant). In order to ensure that iridial structures, such as Wolfflin nodules, could be accurately characterized (even in dark irides), a second photograph was taken using 12-Megapixel Fujifilm Finepix S3Pro SLR camera equipped with a 105 mm macro lens, illuminator kit, and infrared filter. These modifications were necessary because certain structures, such as Wolfflin nodules, cannot be visualized in brown eyes because they rest below the melanin layer. Infrared photography is unique because it is primarily sensitive to light with a wavelength of 700-900 nanometers. The pigment layer does not absorb or reflect light in this range, which means that it is effectively transparent to infrared light. In contrast, the collagen fibers and connective tissues underneath the melanin layer do appear to absorb and reflect radiation in that range. This means that photographs taken using infrared technology will be able to pick up information about the structures, such as Wolfflin nodules, that lie beneath the melanin layer.

DNA collection and extraction. 2mL of saliva were collected using the OraGene OG-500 DNA Collection kit. DNA extraction followed recommended procedures recommended by the manufacturer, with an additional cleanup step to increase yield and purity. DNA samples were quantified via Qubit assay. 500ng of total DNA were shipped to the Clinical Genomics Centre at Mt. Sinai Hospital, Toronto, CA for genotyping.

DNA genotyping. The European, East Asian, South Asian, African American, and US Hispanic samples were genotyped on the Illumina Infinium Multi-Ethnic Array at the Clinical Genomics Centre (Mount Sinai Hospital, Toronto, Ontario, Canada). This array includes ~1.7 million markers and was designed to target common genome variants across diverse population groups. Basic QC steps were carried out using the program Genome Studio, following recommendations by Illumina (resulting in ~ 1.4 million markers remaining). Additional QC steps included: removal of samples with missing call

rates < 0.9, removal of samples with sex discrepancies, removal of samples that were outliers for heterozygosity, and removal of related individuals, removal of markers with genotype call rate < 0.95, removal of markers with Hardy Weinberg p values < 10^{-6} , removal of indels, removal of markers with MAF < 0.01, removal of markers not found on 1000 Genomes reference panel or that show other discrepancies regarding chromosome, position, or alleles. The Mexican samples from Palenque were genotyped using the Axiom Biobanking array and underwent standard QC procedures. The Brazilian samples are still in the process of being genotyped for the Illumina MEGA array. QC of those data will follow what is outlined above.

Statistical analyses/methods.

Phenotypic analyses. The three skin M index values of the right upper inner arm were averaged to produce a mean skin M value for each individual. The same process was followed for CIELab measurements to produce mean L^* , a^* , and b^* values for each individual. Pigmentation values were compared among population samples, based on self-reported ancestry. The mean and standard deviation of M and CIELab values were calculated for each population sample. Mean differences among populations for M index and CIELab values were calculated using ANOVA and MANOVA, respectively. M and CIELab values were transformed as needed to meet assumptions of normality prior to analyses.

Hair M and CIELab values were analyzed in the same manner as skin values. When data was available, these values were also compared to self-reported assessments of five different categories of hair color (light blond/red, blond, dark blond/light brown, dark brown, and black). Comparisons of M and CIELab values among these groups were conducted using ANOVA. M and CIELab values are transformed as needed to meet assumptions of normality prior to analyses.

Iris pigmentation and iridial structures were scored using a custom program developed in the laboratory of the Co-PI (<http://iris.davidcha.ca>) and described in Edwards et al. 2016 (Edwards et al. 2016). Briefly, this program was used to crop the pupil and sclera from iris images and then extract a wedge from the iris itself. The program estimated color scores in CIELab color space from pupillary and ciliary zones, providing mean estimates of L^* , a^* , and b^* for each eye. Mean L^* , a^* , and b^* values were calculated for each population and compared using MANOVA. These values can also be compared to the categorical terms used by participants to describe their eye color (these data are obtained from questionnaire answers). These comparisons have not yet been made, but will be completed as part of the broader investigation into genetic variants associated with iris pigmentation. CIELab values were transformed as needed to meet assumptions of normality prior to analyses.

Genomic analyses. Genotypes are phased using the program SHAPEIT2 (<http://www.shapit.fr/>), using the 1000 Genomes populations as reference haplotypes. Genotype data will be pruned for loci showing short-range and long-range linkage disequilibrium ($r^2 > 0.2$) or genomic complexity. The program EIGENSOFT will be used to perform a Principal Component Analysis (PCA) to identify population substructure in each of our samples. Statistically significant PC's will be included as covariates in subsequent association analyses. This is especially important in the admixed population samples, as PC scores are highly correlated with admixture proportions of major parental groups involved in the admixture process (Parra et al. 2011). The program

SNPTEST (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html) is used to test for associations between genotype and phenotype, conditioning on covariates of sex and significant PC components. In genomic regions where multiple significant signals are detected, SNPTEST can also test for associations while conditioning upon the most significant SNP. This will help to distinguish between signals due to linkage with that SNP or independent signals in the region. In admixed samples admixture mapping methods may also be utilized to identify SNPs with significant effects on pigmentary traits.

Population	Skin	Hair	Iris
African American	Quantitative (M, L*, a*, b*), Fitzpatrick self-assessment	Quantitative (M, L*, a*, b*), categorical self-report	Quantitative (L*,a*,b*), categorical self-report
US Hispanic	Quantitative (M, L*, a*, b*), Fitzpatrick self-assessment	Quantitative (M, L*, a*, b*), categorical self-report	Quantitative (L*,a*,b*), categorical self-report
Mexican	Quantitative (M, L*, a*, b*)	Quantitative (M, L*, a*, b*)	NA
Brazilian	Quantitative (L*, a*, b*)	Categorical self-report	Categorical self-report
European	Quantitative (M, L*, a*, b*), Fitzpatrick self-assessment	Quantitative (M, L*, a*, b*), categorical self-report	Quantitative (L*,a*,b*), categorical self-report
East Asian	Quantitative (M, L*, a*, b*), Fitzpatrick self-assessment	Quantitative (M, L*, a*, b*), categorical self-report	Quantitative (L*,a*,b*), categorical self-report
South Asian	Quantitative (M, L*, a*, b*), Fitzpatrick self-assessment	Quantitative (M, L*, a*, b*), categorical self-report	Quantitative (L*,a*,b*), categorical self-report

Table 1: Phenotypic data available from each population sample.

Population Sample	DNA collected	DNA extracted	Genotyping complete	QC	Phasing	Population Stratification	Association testing
African American	X	X	X	X	In process	Pending	Pending
US Hispanic	X	X	X	X	In process	Pending	Pending
Mexican	X	X	X	X	In process	Pending	Pending
Brazilian	X	X	In process	Pending	Pending	Pending	Pending
European	X	X	X	X	X	X	In process
East Asian	X	X	X	X	X	X	complete
South Asian	X	X	X	X	X	X	In process

Table 2: State of genomic analysis for population samples.

Results

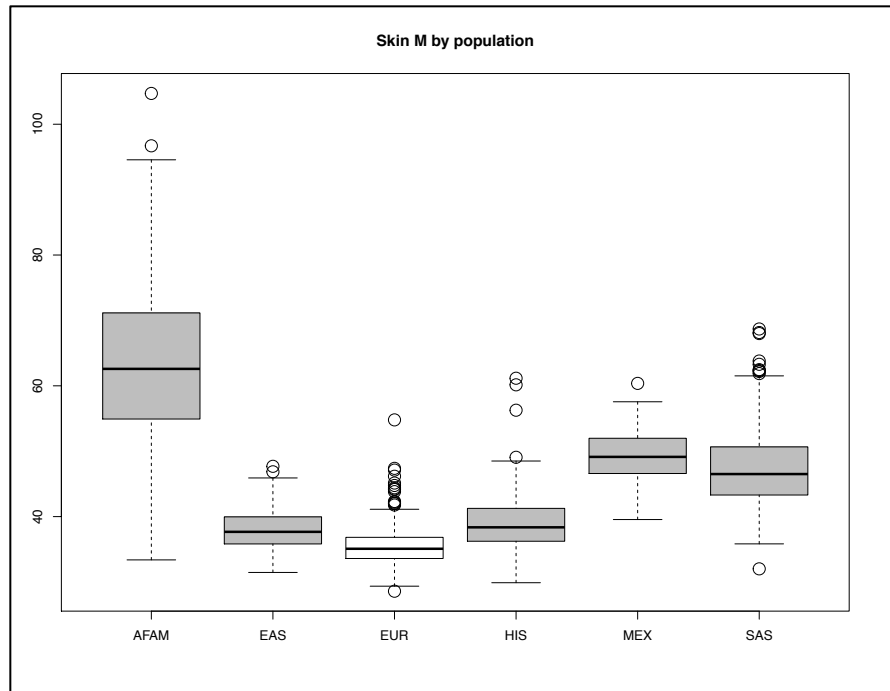
Skin pigmentation. The mean and standard deviation of MI and CIELab values were calculated for all populations (Table 3). As expected, Europeans exhibit the lowest MI

(and highest L^* values), while African Americans and South Asians exhibit the highest MI (and lowest L^* values). It is also notable that skin pigmentation is the least variable in the European sample, but shows a much wider range of variation in the admixed (African American, US Hispanic, Mexican) and South Asian samples (Figure 1). Mean skin MI is significantly different among all six populations ($F = 1044$, $df = 5$, $p < 2 \times 10^{-16}$). A pairwise t-test (with Bonferroni correction) comparing skin MI means among all pairs of populations indicates that all populations are significantly different from each other ($p < 0.00001$) with the exception of the US Hispanic-East Asian comparison ($p = 0.51$) and the Mexican-South Asian comparison ($p = 0.10$). This suggests that these differences are not driven by the darker skin color of the African American sample alone. A MANOVA comparing CIELab skin values confirmed that all six populations are significantly different from each other in CIELab space as well.

Population	N	MI (S.D.)	L^* (S.D.)	a^* (S.D.)	b^* (S.D.)
African American	259	63.1 (12.5)	17.24 (6.5)	12.7 (2.1)	13.63 (4.0)
US Hispanic	90	39.2 (5.2)	33.3 (5.8)	13.5 (1.8)	17.12 (1.7)
Mexican	101	49.0 (4.2)	24.8 (3.8)	15.0 (1.6)	15.2 (1.0)
European	583	35.3 (2.8)	38.0 (3.2)	12.4 (2.0)	14.0 (1.9)
East Asian	425	37.9 (2.9)	35.2 (3.3)	12.8 (1.8)	15.8 (1.7)
South Asian	314	47.5 (6.1)	26.8 (4.7)	13.6 (1.3)	15.2 (1.5)

Table 3: Mean (and standard deviation) for skin MI and CIELab values for the study population samples.

Hair. The mean and standard deviations for hair MI and CIELab values can be found in table 4. Europeans are once again notable for their lower MI values. In contrast to skin MI, Europeans display the widest variation in hair MI (Figure 2). Europeans also exhibit the highest and most variable values for CIELab measurements (Figures 2-4), most likely reflecting the greater diversity of hair colors, including lighter hair, in European populations (this was evident in MI and CIELab values as well as through self-report data obtained from questionnaires). Mean MI differed significantly among the populations ($F = 193.9$, $df = 5$, $p < 2 \times 10^{-16}$). A pairwise t-test (with Bonferroni correction) comparing hair MI means among all pairs of populations indicates that this is driven largely by differences between Europeans and other populations (all comparisons significant at $p < 2 \times 10^{-16}$) and between the US Hispanic sample and other populations ($p < 0.001$). Differences between US Hispanics and other groups may be explained by the lighter hair colors of this group (Table 4, Figure 2). A MANOVA of hair CIELab values confirmed that the populations are significantly different from each other across all the L^* , a^* , and b^* dimensions of CIELab space ($p < 2 \times 10^{-16}$).



Hair color is often reported (in witness statements, drivers licenses, etc.) categorically, rather than in quantitative terms. We evaluated the relationship between self-reported hair color and the quantitative L^* , a^* , and b^* values in the European data sample (the sample with the most variation in hair color). Using information from the participant questionnaires, we divided hair color into five categories: light blond or red, blond, dark blond or light brown, dark brown, or black. Distributions of L^* , a^* , and b^* values for these five categories are shown in Figures 5-7. Significant differences in all three values were observed across the three categories ($p < 0.001$). Higher values of L^* , a^* , and b^* are more common in lighter hair colors. Darker hair categories showed a narrower range of CIELAB values than lighter colors, with those in the light blond/red hair category exhibiting the widest distribution.

Population	N	M (S.D.)	L^* (S.D.)	a^* (S.D.)	b^* (S.D.)
African American	259	167.1 (12.5)	1.32 (6.50)	0.20 (2.12)	0.82 (4.00)
US Hispanic	90	149.8 (17.3)	2.06 (1.2)	0.49 (0.99)	1.35 (1.21)
Mexican	101	163.7 (16.7)	1.53 (0.60)	0.11 (0.38)	0.83 (0.40)
European	583	129.4 (24.4)	3.522 (2.870)	1.584 (1.752)	2.424 (2.322)
East Asian	425	162.8 (18.9)	1.559 (0.788)	0.143 (0.229)	0.747 (0.511)
South Asian	314	164.9 (20.0)	1.558 (1.176)	0.231 (0.547)	0.766 (0.682)

Table 4: Mean (and standard deviation) for hair MI and CIELab values for the study population samples.

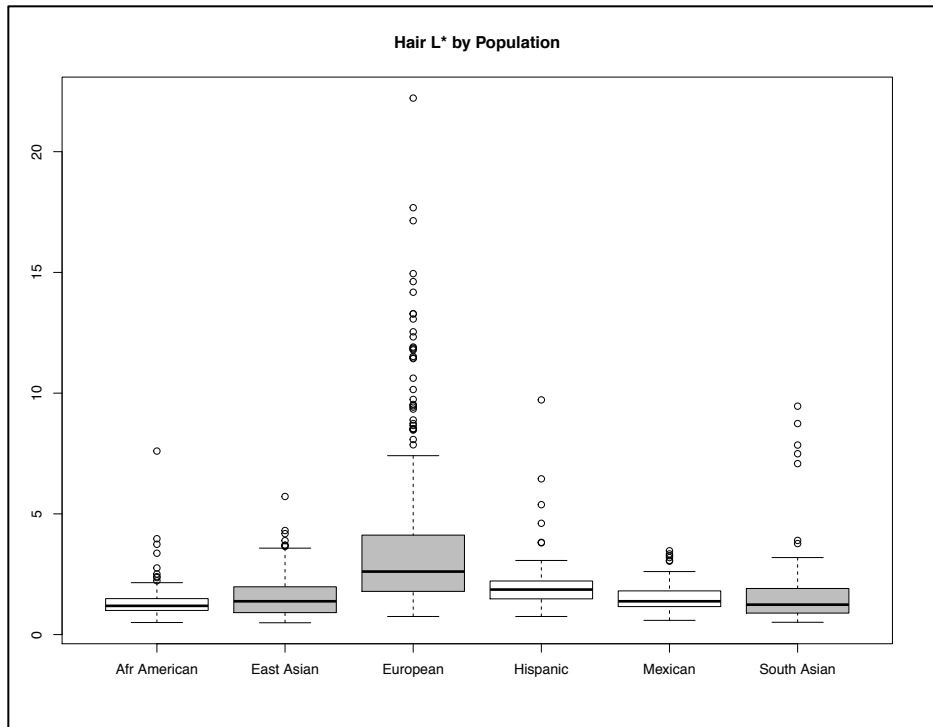


Figure 2: Distribution of hair L*(y axis) values by population.

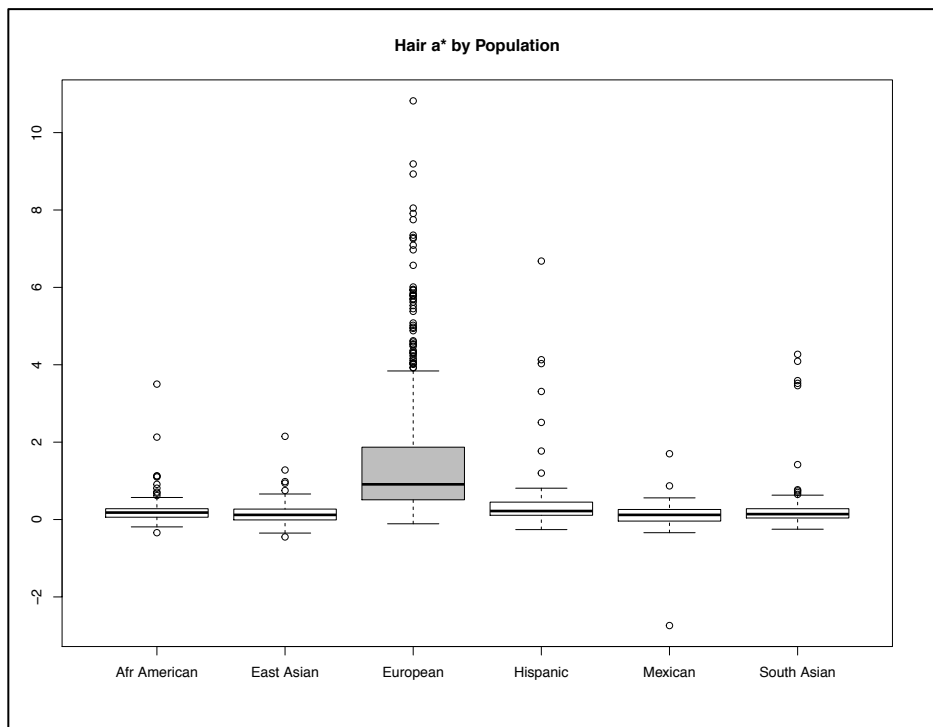


Figure 3: Distribution of hair a* values (y-axis) by population.

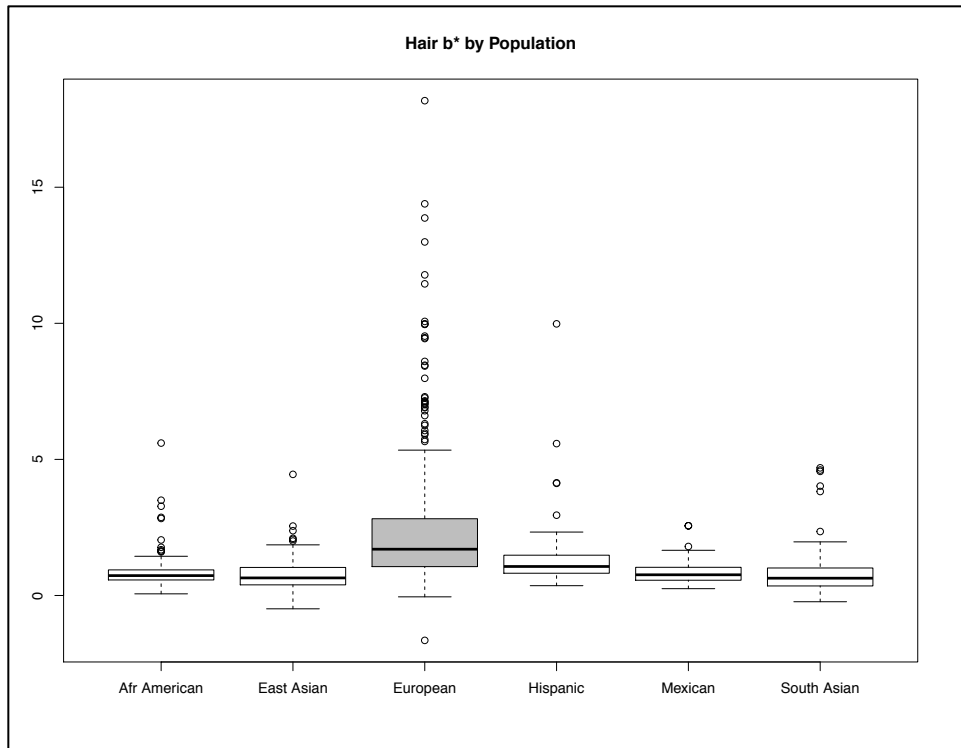


Figure 4: Distribution of hair b* values (y-axis) by population.

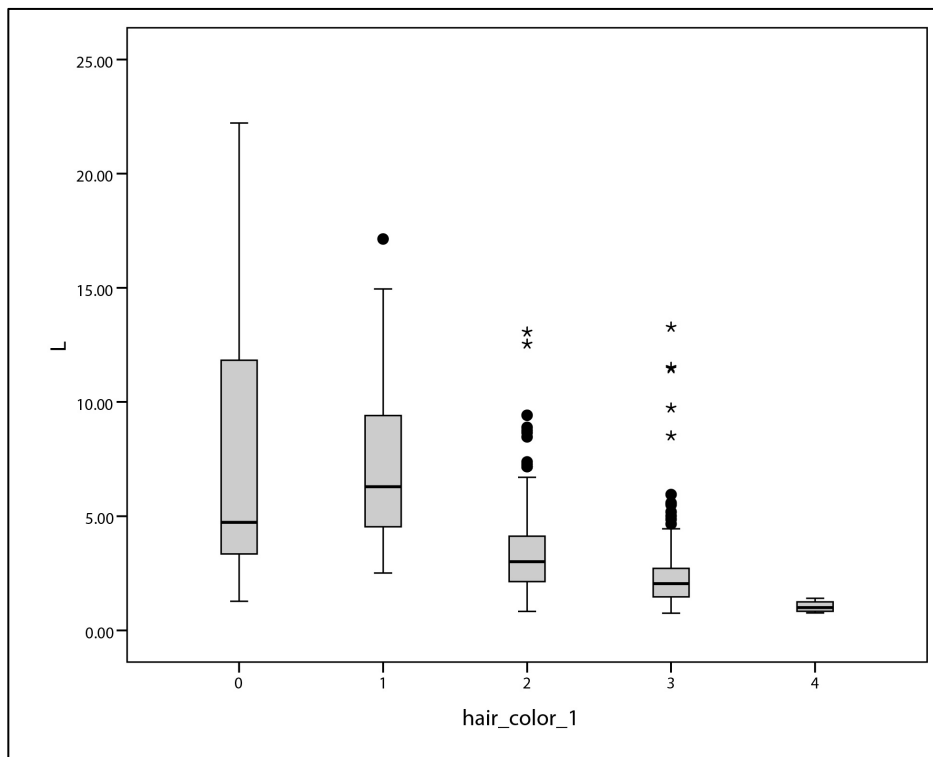


Figure 5: Distribution of L* values across each of five self-described hair color categories in Europeans: 0: Light blond/red; 1: blond; 2: light brown/dark blond; 3: dark brown; 4: black.

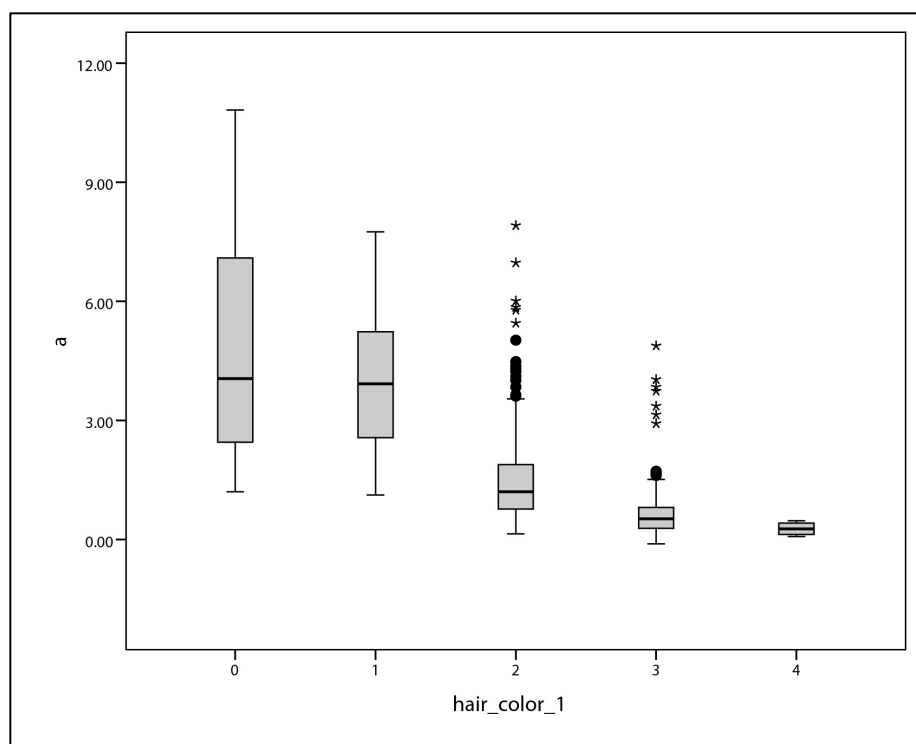


Figure 6: Distribution of a^* values across each of five self-described hair color categories in Europeans: 0: Light blond/red; 1: blond; 2: light brown/dark blond; 3: dark brown; 4: black.

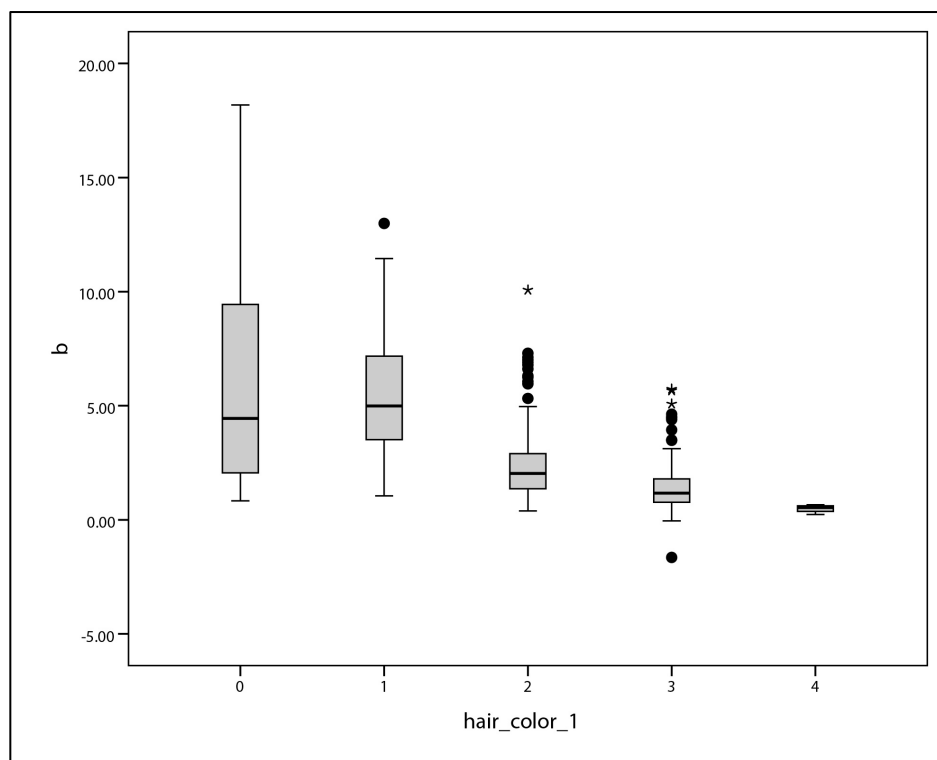


Figure 7: Distribution of b^* values across each of five self-described hair color categories in Europeans: 0: Light blond/red; 1: blond; 2: light brown/dark blond; 3: dark brown; 4: black.

Iris pigmentation. Summaries of variation in iris pigmentation for the European, East Asian, and South Asian population samples can be found in Edwards (Edwards et al. 2016). Briefly, that study demonstrated eyes that were self-described as “blue” tended to exhibit high L^* values and negative a^* and b^* values. At the opposite end of the spectrum, self-reported brown eyes exhibit low L^* and high a^* and b^* values. The authors were also able to draw some distinction between light and dark brown eyes, with the latter typically having lower values for L^* , a^* , and b^* . Mean and ranges of L^* , a^* , and b^* variation for the African American and US Hispanic population samples are shown in Table 5. Compared to the European population samples reported by Edwards and colleagues., both the African American and US Hispanic samples tend to have darker eye colors (evidenced by lower mean values of L^*). However, values in both admixed populations are highly variable. Figure 8 plots the distribution of these iris values across CIELab color space for the African American and US Hispanic samples. These plots show a range of variation in iris pigmentation in both groups, with most eye colors ranging from dark to light brown. It is notable, though that some individuals from both the African American and US Hispanic samples exhibit relatively light eye colors (including one US Hispanic individual with CIELab values indicative of blue eye color.

Population	L^* (Min-Max)	a^* (Min-Max)	b^* (Min-Max)
African American	8.54 (2.30-44.19)	3.37 (-2.22-17.25)	2.26 (-2.23-21.92)
US Hispanic	18.46 (4.366-76.43)	6.06 (-3.25-17.95)	6.45 (-11.70-20.68)

Table 5: Mean (and minimum-maximum) values of iris L^* , a^* , and b^* values for the African American and US Hispanic population samples.

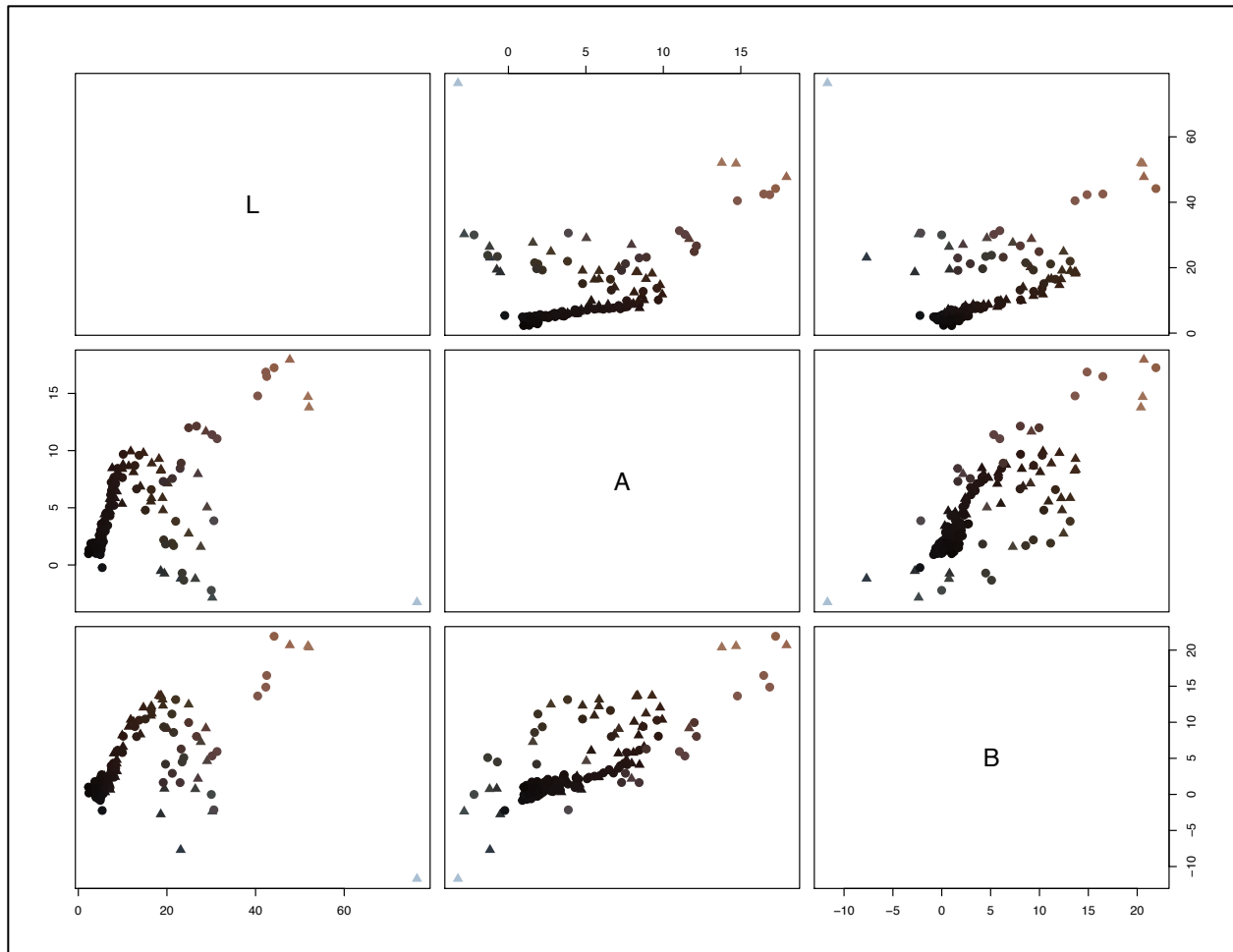


Figure 8: Scatterplot matrix of CIELab iris values for African American and US Hispanic individuals. The color of each point corresponds to iris color in CIELab space. African American individuals are indicated by circles, US Hispanic individuals by triangles.

Genetic associations. Because analyses of genomic data in the admixed populations are still ongoing, reports of associations at this time are limited. However, in a preliminary analysis a small number of SNPs previously associated with pigmentation phenotype were genotyped in the European sample and reported in Norton et al. 2016 (Norton et al. 2016). One of these (rs1426654 in the gene *SLC24A5*) was excluded from further analyses because the allele frequency of the ancestral G allele (associated with darker skin color) was < 1%. Associations between the remaining eight SNPs and skin MI and hair CIELab values were tested for using an additive model (Tables 6 and 7). The derived alleles at three of these SNPs (rs16891982, rs12203592, and rs12913832) were significantly associated with lighter skin color in this European sample. Derived alleles in five SNPs were significantly associated with lighter hair pigmentation (higher L*, a*, or b* values): rs16891982, rs1291832, rs7495174, rs4778138, and rs12896399. The derived allele (T) at rs12203592 was associated with darker skin color. Tests that included rs12913832 (shown to have a strong effect on hair color) as a covariate indicated that significant associations between rs7495174 and rs4778138 (also located

in the *OCA2/HERC2*) region were likely driven by linkage disequilibrium with rs12913832.

Skin M				
Gene	Polymorphism	Effect allele	BETA	p-value
SLC45A2	rs16891982	G	-0.0431	1.21E-07
IRF4	rs12203592	T	-0.0286	2.71E-07
<i>TYR</i>	rs1393350	A	-0.0104	0.0415
<i>SLC24A4</i>	rs12896399	T	-0.0005	0.9086
<i>OCA2</i>	rs1800407	A	-0.0047	0.6258
<i>OCA2</i>	rs4778138	A	-0.0126	0.0328
<i>OCA2</i>	rs7495174	A	-0.0207	0.0091
HERC2	rs12913832	G	-0.0178	9.95E-05

Table 6: Linear regression coefficients and p-values for each of the 8 SNPs in the European sample for skin *M* index. Significant results (after a Bonferroni correction) are highlighted in **bold**.

Gene	Polymorphism	Effect allele	BETA	p-value
Hair L*				
<i>SLC45A2</i>	rs16891982	G	0.2270	0.0075
IRF4	rs12203592	T	-0.1929	0.0007
<i>TYR</i>	rs1393350	A	0.0441	0.3973
SLC24A4	rs12896399	T	0.1246	0.0046
<i>OCA2</i>	rs1800407	A	0.0972	0.3024
<i>OCA2</i>	rs4778138	A	0.1297	0.0295
OCA2	rs7495174	A	0.2410	0.0024
HERC2	rs12913832	G	0.2280	7.21E-07
Hair a*				
SLC45A2	rs16891982	G	0.2522	0.0056
IRF4	rs12203592	T	-0.2454	5.55E-05
<i>TYR</i>	rs1393350	A	0.0062	0.9121
SLC24A4	rs12896399	T	0.1632	0.0005
<i>OCA2</i>	rs1800407	A	0.0882	0.3833
OCA2	rs4778138	A	0.2088	0.0011
OCA2	rs7495174	A	0.2750	0.0013
HERC2	rs12913832	G	0.3271	2.08E-11
Hair b*				
SLC45A2	rs16891982	G	0.1575	0.0050
IRF4	rs12203592	T	-0.1426	0.0050
<i>TYR</i>	rs1393350	A	0.0073	0.8312
SLC24A4	rs12896399	T	0.0923	0.0015
<i>OCA2</i>	rs1800407	A	0.0610	0.3270
<i>OCA2</i>	rs4778138	A	0.0987	0.0123
OCA2	rs7495174	A	0.1658	0.0016

HERC2	rs12913832	G	0.1583	1.90E-07
--------------	-------------------	----------	---------------	-----------------

Table 7: Linear regression coefficients and p-values for each of the 8 SNPs in the European sample for hair CIELab values. Significant results (after a Bonferroni correction) are highlighted in **bold**.

A GWAS of skin and iris pigmentation variation in the East Asian population sample has been completed—full results can be read in Rawofi et al. (Rawofi et al. 2017). An important feature of this work (also emphasized in Edwards et al. 2016), is the distribution of L*, a*, and b* iris values. While eye color in this population would traditionally be classified as “brown”, there is notable variation across CIELAB space (indicating variation in shades of brown) demonstrating that the use of broad categories to describe iris phenotype may mask important and relevant phenotypic variation.

No associations reaching genome-wide significance ($p = 5 \times 10^{-8}$) were observed for skin M index (Figure 9). Association tests between five common markers (MAF > 1%) with the suggestive p-values ($< 10^{-5}$) were carried out in two replication samples. An association between one of these, rs2373391 (located in the *ZNF804B* gene on chromosome 7), was replicated in one of these cohorts, with the A allele being significantly associated with lighter skin color. A single SNP, rs76930569 in the *OCA2* region, was significantly associated with two dimensions of iris color (Figure 10). This association is likely driven by LD between this SNP and rs1800414, a nonsynonymous polymorphism previously associated with skin pigmentation. This SNP was estimated to explain 11.9%, 10.4%, and 6% of variation in iris b*, a*, and L* values in this sample.

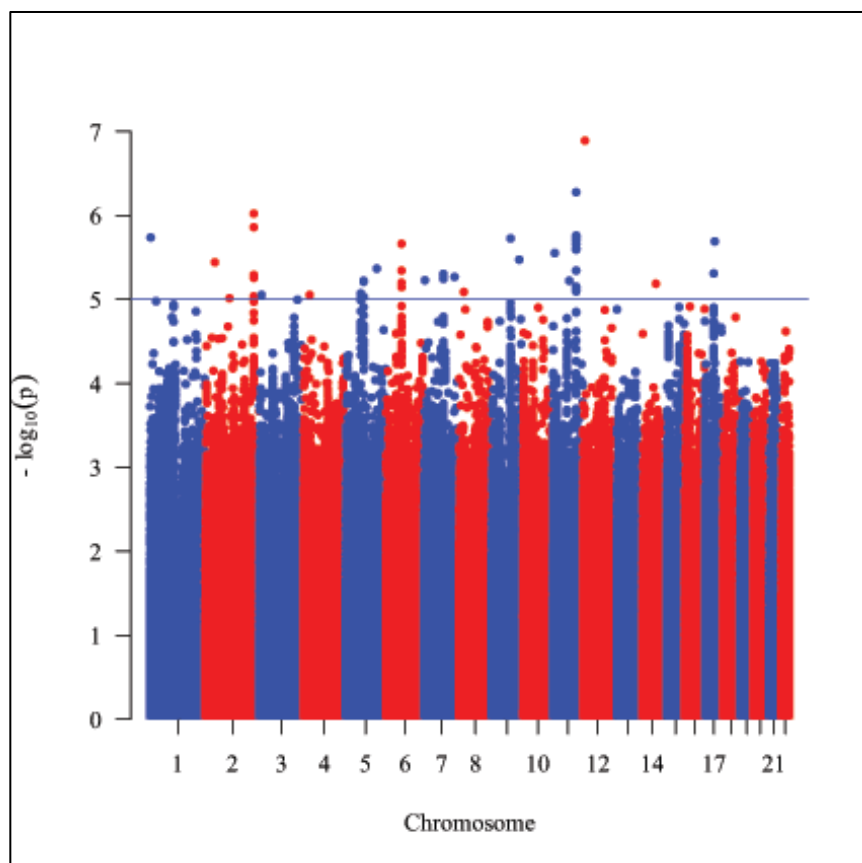


Figure 9: Manhattan plot reporting association results for skin pigmentation in East Asian sample (from Rawofi et al. 2017)

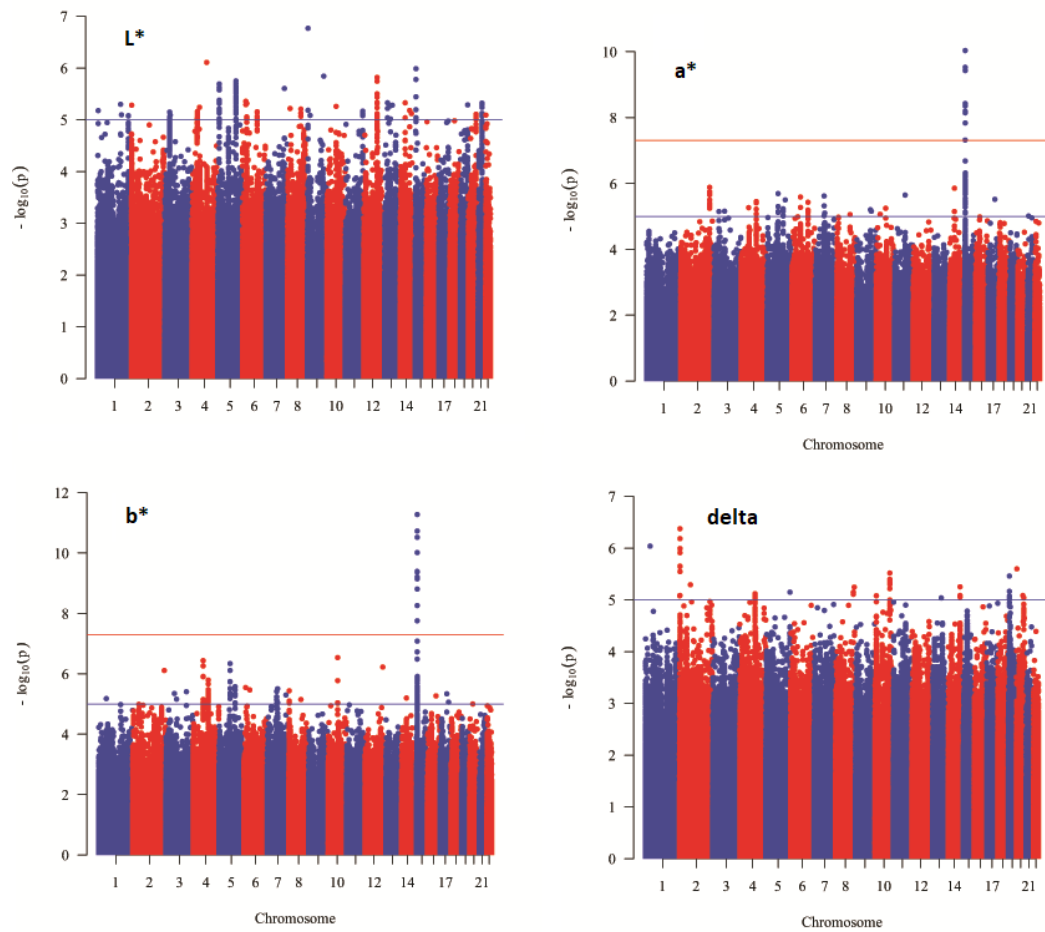


Figure 10: Manhattan plot reporting associations for iris pigmentation (L^* , a^* , and b^*), as well as for delta (the difference the ciliary and pupillary regions of the iris). From Rawofi et al. 2017.

Conclusions

Discussion. As analyses of genomic data from the admixed populations in this study is still ongoing this discussion will focus largely on the phenotyping data collected here. One of the key features of this study compared to others is the use of quantitative methods to assess skin, hair and iris color. This was a specific and deliberate choice on the part of the PI and co-PI for a number of reasons. First, our experience with these methods in previous studies has emphasized to us the continuous nature of these pigimentary traits, and the degree to which phenotypic variation can be masked under simple categorical labels. In addition, our previous work also highlighted to us that the use of categorical labels to describe these phenotypes led to the mistaken belief that there was little variation in traits such as hair and iris color outside of European populations. Second, previous work examining the genetic basis of iris pigmentation (Liu et al. 2010) has shown that the use of quantitative metrics to assess iris pigmentation resulted in improved power to detect associated genetic loci.

Skin pigmentation. We observe significant variation in skin pigmentation (measured as MI or with CIELab values) between the populations included here. Pairwise comparisons among populations indicate that these differences are not driven by one or two outlier populations. However, we also observed that the distributions of quantitative skin pigmentation measurements overlapped for all six populations (Figure 1), including the two populations with the most discordant mean skin MI values (Europeans and African Americans). Despite the additional effort and technical equipment required to obtain quantitative pigmentation estimates, we feel that the benefits (better understanding of within and between population, improved power to detect genetic loci) outweigh these difficulties.

As expected, we observed significant variation in skin pigmentation in the admixed population samples, consistent with studies of other admixed populations (Andrade et al. 2017; Beleza et al. 2013; Bonilla et al. 2004a; Bonilla et al. 2004b; Hernandez-Pacheco et al. 2017; Martin et al. 2017; Shriver et al. 2003). This is important, as it reflects not only variation in ancestry proportions as well as possible differences in the genetic architecture of skin pigmentation in these populations due to distinct combinations of pigimentary alleles (Hernandez-Pacheco et al. 2017). This variation also highlights the fact that our admixed populations are drawn from heterogeneous groups. For example, the individuals in the US Hispanic sample report ancestry from Puerto Rico, Mexico, Columbia, Cuba, and Venezuela as well as generally “Latino”. The variation in this sample, as well as variation when compared to the Mexican sample from Palenque is a strong reminder that social labels such as race or ethnicity or geopolitical affiliations do not always map neatly onto phenotype. A similar conclusion can be drawn from the South Asian sample, which consisted of participants claiming ancestry from Bangladesh, India, Pakistan, and Sri Lanka, and also showed a wide range of variation in skin M index. This heterogeneity also emphasizes the importance of being able to control for population substructure in the GWAS analyses, as failure to do so may result in false-positive associations related to differences in ancestry proportions.

We observed the greatest variation in hair pigmentation in the European sample, where hair color was generally lighter (higher values of L^* , a^* , and b^*) and more variable

than in other populations. However, differences in hair color between other populations was also observed, partly due to the wide range of variation observed in the US Hispanic sample. Perhaps surprisingly, differences in L^* , a^* , and b^* values were also observed between the East and South Asian samples, where hair would typically be classified as brown or black. These differences were largely attributed to higher a^* values in the South Asian sample. Notably no significant differences in MI were observed between these two groups, suggesting that CIELAB space may allow for more nuanced phenotyping of hair color than MI or categorical labels. This also suggests that it may be possible to identify loci with subtle effects on hair color variation (particularly in the a^* component) across populations.

We compared quantitative estimates of hair color in the European sample to self-reported hair color in order to assess whether such self-categorization could accurately capture phenotypic variation. Our results (Figures 5, 6, and 7) demonstrate why using self-reported categories in association studies is likely to be a suboptimal strategy. The broad distribution of CIELAB values within each category (particularly among lighter hair colors) suggests that a substantial amount of nuanced phenotypic information is lost when only self-reported categories are used. It is also clear that there is overlap in CIELAB values across categories, meaning that two individuals who self-report different hair colors may have similar hair reflectance values. This may impact the effects of downstream association analyses.

Our quantitative characterization of iris pigmentation confirmed previous reports demonstrating that eye color better fits a continuous, rather than categorical distribution (Beleza et al. 2013; Edwards et al. 2016; Edwards et al. 2011; Liu et al. 2010). Details of the distribution of L^* , a^* , and b^* values for the European, East, and South Asian samples can be found in Edwards et al. 2016 and a preliminary comparison of a subset of those and the African American and US Hispanic samples is reported in Norton et al. 2016a. One of the key conclusions drawn from the admixed populations we examined is that while eye color in many of these individuals would classically be categorized as “brown” there is substantial variation in CIELab space. The African American and US Hispanic samples observed here exhibit greater variation in L^* and a^* values than the East and South Asian samples, although variation in the b^* component is roughly similar. Several of these admixed individuals exhibit values consistent with intermediate iris color phenotypes, and one US Hispanic individual exhibits blue eye color (Figure 8). Once again, this highlights the wide range of variation present in admixed populations and cautions against assuming simplistic correlations between pigmentary phenotypes and social categories such as race or ethnicity.

Our preliminary analyses testing for associations between nine known pigmentary alleles and variation in the European sample confirmed previously identified associations with skin color for three SNPs located in the genes *SLC45A2*, *IRF4*, and *HERC2*. We also confirmed the association of four genes (*HERC2*, *IRF4*, *SLC24A4*, and *SLC45A2*). We noted that the derived T allele at the rs12203592 polymorphism was associated with lighter skin but darker hair color in our European sample (Norton et al. 2016). This suggests that the quantitative phenotyping methods that we used facilitated the detection of alleles having different effects in different tissues. These more subtle effects may be masked in non-European populations, especially when these complex phenotypes are considered in only categorical terms.

Edwards et al. (2016) tested for associations between iris L*, a*, and b* values and 14 pigmentation-related SNPs in the European, East Asian, and South Asian samples included in this study. They confirmed the association of five of those SNPs (all of which are presently on the IrisPlex forensic panel) with iris pigmentation in the European sample. They found that the rs12913892 polymorphism in *HERC2* played an important role in modulating iris color in the South Asian sample, primarily through the modulation of brown eye color (dark vs. light). While having two copies of the derived allele at this locus is associated with a decrease in b* values in Europeans, no such decrease was observed in the South Asian sample. The authors postulate that rs12913892 may impact iris pigmentation differently in European and South Asian populations because of its interactions with other (as yet unknown) loci that are common in European, but not South Asian populations. These results suggest that the genetic architecture of iris pigmentation may vary among populations, a critical factor when considering the development and broad application of marker panels to predict iris phenotype across diverse populations.

A GWAS of skin and iris pigmentation in the East Asian sample is complete (Rawofi et al. 2017). This study did not identify any SNPs explaining variation in skin pigmentation variation that reached genome-wide significance, despite having good power (>0.8) to detect loci explaining ~12.5% of variation in skin pigmentation phenotype. This suggests that there are no variants segregating in the sample that have large effects on skin pigmentation, indicating that the genetic architecture of skin pigmentation variation in East Asians may be regulated by many markers of relatively small effect. One SNP in the gene *ZNF804B* reached suggestive significance, and that will merit follow-up in future studies. Notably, the derived A allele at this SNP (rs2373391), which is associated with lighter skin pigmentation, is polymorphic in admixed populations of the 1000 Genomes Panel (with frequencies ranging from 0.234-0.429), suggesting that we should be able to evaluate its effects on pigmentation in our admixed samples. Finally, by using quantitative phenotyping methods this study was able to identify a variant, rs1800414 in the *OCA2* gene that influenced iris pigmentation. However, this SNP could only explain ~10% of variation in observed L*, a*, and b* values, leaving a substantial amount of variation unexplained.

Implications for policy and practice. Here we have emphasized the use of quantitative methods to characterize pigmentation variation, and have highlighted the importance of such methods in the identification of genes that can explain moderate effects on phenotype, particularly in non-European populations. We also believe that presenting these phenotypes as continuous, rather than discrete, traits is critical to prevent individuals from incorrectly perceiving socially-defined racial or ethnic groups as being homogeneous for a given phenotype and/or from being biologically distinct from other such groups. In other words, we caution strongly against interpreting a phenotypic prediction of skin, hair, or iris color as evidence of a particular socially-constructed identity. Supporting this, we point to the overlap in skin, hair, and iris MI and CIELab values observed across populations in this study. We also observed a wide range of variation in pigmentation phenotypes in our study populations, even for phenotypes (e.g. iris color) believed to be generally homogeneous.

If FDP methods can be refined such that predictions can be made about pigmentation phenotypes (e.g., skin pigmentation falls within a certain range of MI

values) it is currently unclear how to translate this information so that it is useful for investigators. While it is generally easy to understand that higher MI values correlate with darker skin colors, it is less clear how easily an individual will be able to understand what a predicted MI value of 40-55 means in terms of physical appearance. One possible way forward to is to assemble suites of photographs of skin representing different MI ranges to be used as a training tool. It is unclear how well the human eye can discriminate among these different ranges, however. In addition, given the fact that hair and eye color are commonly reported in categorical terms by eye witnesses and on official documents (e.g. drivers' licenses) it may be useful to better understand how those descriptive categories correlate with quantitative pigmentation assessments.

Several studies have published panels of markers that may be used to predict skin, hair, and/or iris pigmentation with various degrees of accuracy (Ruiz et al. 2013; Walsh et al. 2017; Walsh et al. 2011b; Walsh et al. 2013). While many of these were developed first in European populations, there is some evidence that these panels may be useful in predicting phenotype in more diverse population samples, including admixed populations (Andrade et al. 2017; Freire-Aradas et al. 2014), although prediction rates for "intermediate" phenotypes seem to lag behind prediction rates for phenotypes at the extremes of distributions (e.g. prediction success is higher for blue/brown eye color vs. intermediate eye color). In some cases prediction models were derived (or tested) using unphenotyped samples, and were deemed successful when their predictions met the authors' assumptions of phenotype in those populations. Our work here and others investigations into pigmentation variation, particularly in African populations, suggests that these may be faulty assumptions, masking substantial pigmentation variation in these populations.

Of particular concern regarding the development of broadly applicable panels of markers to be used in FDP is that loci commonly included on such panels may not accurately predict phenotype with suitable accuracy in certain individuals. Edwards (Edwards et al. 2016) work noting the different expression of rs1291382 in European and South Asian populations provides one such example where this concern seems justified. There is also growing evidence that the genetic architecture of pigmentation variation may differ substantially across populations. Much of our knowledge of pigmentation genetics comes from studies in European populations, where it seems like variation may be controlled by relatively small numbers of loci of large effect. The East Asian pigmentation GWAS reported by Rawof and colleagues suggests that this may not be the case in East Asian populations (Rawof et al. 2017). In addition, recent work identifying novel loci with small but significant effects on pigmentation in African and admixed populations (Crawford et al. 2017; Hernandez-Pacheco et al. 2017; Martin et al. 2017) all suggest that accurate phenotypic prediction may be substantially more challenging than anticipated. It may be that while loci underlying the genetic architecture of pigmentation across diverse samples can be identified, the number of loci required for accurate prediction may not be suitable in forensic settings.

Future research. The completion of the genotypic analyses in our admixed samples will clearly help to improve our understanding of the genetic architecture of pigmentation variation in admixed populations, and may facilitate the discovery of loci that can be used to predict pigmentation across a broader array of populations. Replication studies in other admixed populations of significant and suggestive signals will be carried out.

As noted above, it is currently unclear how quantitative ranges of pigmentation values can be translated into information that can be used productively by investigators on the ground. It is also unclear how providing information about pigmentation phenotype might influence an individual's perception of the race or ethnicity of an unknown individual. Investigations exploring how well quantitative measures of pigmentation map onto commonly used hair and eye color categories (as we did for hair color in the European sample) may be useful here, as would investigations into how well people are able to discriminate among different MI ranges.

References

- Adhikari K, Mendoza-Revilla J, Sohail A, Fuentes-Guajardo M, Lampert J, Chacon-Duque JC, Hurtado M, Villegas V, Granja V, Acuna-Alonzo V et al. . 2019. A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nat Commun* 10(1):358.
- Akey JM, Wang H, Xiong M, Wu H, Liu W, Shriver MD, 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(6):516-520.
- Andrade ES, Fracasso NCA, Strazza Junior PS, Simoes AL, and Mendes-Junior CT. 2017. Associations of OCA2-HERC2 SNPs and haplotypes with human pigmentation characteristics in the Brazilian population. *Leg Med (Tokyo)* 24:78-83.
- Beleza S, Johnson NA, Candille SI, Absher DM, Coram MA, Lopes J, Campos J, Araujo, II, Anderson TM, Vilhjalmsen BJ et al. . 2013. Genetic architecture of skin and eye color in an African-European admixed population. *PLoS Genet* 9(3):e1003372.
- Bonilla C, Boxill LA, Donald SA, Williams T, Sylvester N, Parra EJ, Dios S, Norton HL, Shriver MD, and Kittles RA. 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(5):402-406.
- Bonilla C, Parra EJ, Pfaff CL, Dios S, Marshall JA, Hamman RF, Ferrell RE, Hoggart CL, McKeigue PM, and Shriver MD. 2004a. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann Hum Genet* 68(Pt 2):139-153.
- Bonilla C, Shriver MD, Parra EJ, Jones A, and Fernandez JR. 2004b. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. *Human genetics* 115(1):57-68.
- Branicki W, Liu F, van Duijn K, Draus-Barini J, Pospiech E, Walsh S, Kupiec T, Wojas-Pelc A, and Kayser M. 2011. Model-based prediction of human hair color using DNA variants. *Hum Genet* 129(4):443-454.
- Crawford NG, Kelly DE, Hansen MEB, Beltrame MH, Fan S, Bowman SL, Jewett E, Ranciaro A, Thompson S, Lo Y et al. . 2017. Loci associated with skin pigmentation identified in African populations. *Science* 358(6365).
- Diffey BL, Oliver RJ, and Farr PM. 1984. A portable instrument for quantifying erythema induced by ultraviolet radiation. *Br J Dermatol* 111(6):663-672.
- Edwards M, Bigham A, Tan J, Li S, Gozdzik A, Ross K, Jin L, and Parra EJ. 2010. Association of the OCA2 polymorphism His615Arg with melanin content in east Asian populations: further evidence of convergent evolution of skin pigmentation. *PLoS Genet* 6(3):e1000867.
- Edwards M, Cha D, Krithika S, Johnson M, Cook G, and Parra EJ. 2016. Iris pigmentation as a quantitative trait: variation in populations of European, East Asian and South Asian ancestry and association with candidate gene polymorphisms. *Pigment Cell Melanoma Res* 29(2):141-162.
- Edwards M, Gozdzik A, Ross K, Miles J, and Parra EJ. 2011. Technical note: Quantitative measures of iris color using high resolution photographs. *American journal of physical anthropology*.

- Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer KW, and Hansen L. 2008. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the *HERC2* gene inhibiting *OCA2* expression. *Human genetics* 123(2):177-187.
- Eilers S, Bach DQ, Gaber R, Blatt H, Guevara Y, Nitsche K, Kundu RV, and Robinson JK. 2013. Accuracy of self-report in assessing Fitzpatrick skin phototypes I through VI. *JAMA Dermatol* 149(11):1289-1294.
- Fitzpatrick TB. 1988. The validity and practicality of sun-reactive skin types I through VI. *Arch Dermatol* 124(6):869-871.
- Freire-Aradas A, Ruiz Y, Phillips C, Maronas O, Sochtig J, Tato AG, Dios JA, de Cal MC, Silbiger VN, Luchessi AD et al. . 2014. Exploring iris colour prediction and ancestry inference in admixed populations of South America. *Forensic Sci Int Genet* 13:3-9.
- Frudakis T, Terravainen T, and Thomas M. 2007. Multilocus *OCA2* genotypes specify human iris colors. *Human genetics* 122(3-4):311-326.
- Frudakis T, Thomas M, Gaskin Z, Venkateswarlu K, Chandra KS, Ginjupalli S, Gunturi S, Natrajan S, Ponnuswamy VK, and Ponnuswamy KN. 2003. Sequences associated with human iris pigmentation. *Genetics* 165(4):2071-2083.
- Grimes EA, Noake PJ, Dixon L, and Urquhart A. 2001. Sequence polymorphism in the human melanocortin 1 receptor gene as an indicator of the red hair phenotype. *Forensic Sci Int* 122(2-3):124-129.
- Han J, Kraft P, Nan H, Guo Q, Chen C, Qureshi A, Hankinson SE, Hu FB, Duffy DL, Zhao ZZ et al. . 2008. A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. *PLoS Genet* 4(5):e1000074.
- Hernandez-Pacheco N, Flores C, Alonso S, Eng C, Mak AC, Hunstman S, Hu D, White MJ, Oh SS, Meade K et al. . 2017. Identification of a novel locus associated with skin colour in African-admixed populations. *Sci Rep* 7:44548.
- Jonnalagadda M, Norton H, Ozarkar S, Kulkarni S, and Ashma R. 2016. Association of genetic variants with skin pigmentation phenotype among populations of west Maharashtra, India. *American Journal of Human Biology* 28(5):610-618.
- Kayser M, and de Knijff P. 2011. Improving human forensics through advances in genetics, genomics and molecular biology. *Nat Rev Genet* 12(3):179-192.
- Kayser M, Liu F, Janssens AC, Rivadeneira F, Lao O, van Duijn K, Vermeulen M, Arp P, Jhamai MM, van Ijcken WF et al. . 2008. Three genome-wide association studies and a linkage analysis identify *HERC2* as a human iris color gene. *American journal of human genetics* 82(2):411-423.
- Kayser M, and Schneider PM. 2009. DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. *Forensic Sci Int Genet* 3(3):154-161.
- Keating B, Bansal AT, Walsh S, Millman J, Newman J, Kidd K, Budowle B, Eisenberg A, Donfack J, Gasparini P et al. . 2013. First all-in-one diagnostic tool for DNA intelligence: genome-wide inference of biogeographic ancestry, appearance, relatedness, and sex with the Identitas v1 Forensic Chip. *Int J Legal Med* 127(3):559-572.
- Liu F, Wollstein A, Hysi PG, Ankra-Badu GA, Spector TD, Park D, Zhu G, Larsson M, Duffy DL, Montgomery GW et al. . 2010. Digital quantification of human eye color highlights genetic association of three new loci. *PLoS Genet* 6(5):e1000934.

- Maronas O, Phillips C, Sochtig J, Gomez-Tato A, Cruz R, Alvarez-Dios J, de Cal MC, Ruiz Y, Fondevila M, Carracedo A et al. . 2014. Development of a forensic skin colour predictive test. *Forensic Sci Int Genet* 13:34-44.
- Martin AR, Lin M, Granka JM, Myrick JW, Liu X, Sockell A, Atkinson EG, Werely CJ, Moller M, Sandhu MS et al. . 2017. An Unexpectedly Complex Architecture for Skin Pigmentation in Africans. *Cell* 171(6):1340-1353 e1314.
- Norton HL, Edwards M, Krithika S, Johnson M, Werren EA, and Parra EJ. 2016. Quantitative assessment of skin, hair, and iris variation in a diverse sample of individuals and associated genetic variation. *American Journal of Physical Anthropology* 160(4):570-581.
- Norton HL, Friedlaender JS, Merriwether DA, Koki G, Mgone CS, and Shriver MD. 2006. Skin and hair pigmentation variation in Island Melanesia. *American journal of physical anthropology* 130(2):254-268.
- Parra EJ. 2007. Human pigmentation variation: evolution, genetic basis, and implications for public health. *American journal of physical anthropology Suppl* 45:85-105.
- Parra EJ, Below JE, Krithika S, Valladares A, Barta JL, Cox NJ, Hanis CL, Wachter N, Garcia-Mena J, Hu P et al. . 2011. Genome-wide association study of type 2 diabetes in a sample from Mexico City and a meta-analysis of a Mexican-American sample from Starr County, Texas. *Diabetologia* 54(8):2038-2046.
- Pichon L, Landrine H, Corral I, Hao Y, Mayer J, and K H. 2010. Measuring skin cancer risk in African Americans: Is the Fitzpatrick Skin Type Classification Scale culturally sensitive? *Ethnicity & Disease* 20:174-179.
- Rawofi L, Edwards M, Krithika S, Le P, Cha D, Yang Z, Ma Y, Wang J, Su B, Jin L et al. . 2017. Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry. *PeerJ* 5:e3951.
- Ruiz Y, Phillips C, Gomez-Tato A, Alvarez-Dios J, Casares de Cal M, Cruz R, Maronas O, Sochtig J, Fondevila M, Rodriguez-Cid MJ et al. . 2013. Further development of forensic eye color predictive tests. *Forensic Sci Int Genet* 7(1):28-40.
- Shriver MD, and Parra EJ. 2000. Comparison of narrow-band reflectance spectroscopy and tristimulus colorimetry for measurements of skin and hair color in persons of different biological ancestry. *American journal of physical anthropology* 112(1):17-27.
- Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N et al. . 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Human genetics* 112(4):387-399.
- Spinney L. 2008. Eyewitness identification: line-ups on trial. *Nature* 453(7194):442-444.
- Stokowski RP, Pant PV, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ et al. . 2007. A genomewide association study of skin pigmentation in a South Asian population. *American journal of human genetics* 81(6):1119-1132.
- Sturm RA, Duffy DL, Zhao ZZ, Leite FP, Stark MS, Hayward NK, Martin NG, and Montgomery GW. 2008. A single SNP in an evolutionary conserved region within intron 86 of the *HERC2* gene determines human blue-brown eye color. *American journal of human genetics* 82(2):424-431.

- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, Steinberg S, Gudjonsson SA, Palsson A, Thorleifsson G et al. . 2008. Two newly identified genetic determinants of pigmentation in Europeans. *Nat Genet* 40(7):835-837.
- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A, Thorleifsson G et al. . 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39(12):1443-1452.
- Valenzuela RK, Henderson MS, Walsh MH, Garrison NA, Kelch JT, Cohen-Barak O, Erickson DT, John Meaney F, Bruce Walsh J, Cheng KC et al. . 2010. Predicting phenotype from genotype: normal pigmentation. *J Forensic Sci* 55(2):315-322.
- Walsh S, Chaitanya L, Breslin K, Muralidharan C, Bronikowska A, Pospiech E, Koller J, Kovatsi L, Wollstein A, Branicki W et al. . 2017. Global skin colour prediction from DNA. *Hum Genet* 136(7):847-863.
- Walsh S, Lindenbergh A, Zuniga SB, Sijen T, de Knijff P, Kayser M, and Ballantyne KN. 2011a. Developmental validation of the IrisPlex system: determination of blue and brown iris colour for forensic intelligence. *Forensic Sci Int Genet* 5(5):464-471.
- Walsh S, Liu F, Ballantyne KN, van Oven M, Lao O, and Kayser M. 2011b. IrisPlex: a sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. *Forensic Sci Int Genet* 5(3):170-180.
- Walsh S, Liu F, Wollstein A, Kovatsi L, Ralf A, Kosiniak-Kamysz A, Branicki W, and Kayser M. 2013. The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. *Forensic Sci Int Genet* 7(1):98-115.
- Walsh S, Wollstein A, Liu F, Chakravarthy U, Rahu M, Seland JH, Soubrane G, Tomazzoli L, Topouzis F, Vingerling JR et al. . 2012. DNA-based eye colour prediction across Europe with the IrisPlex system. *Forensic Sci Int Genet* 6(3):330-340.
- Wells GL, Malpass RS, Lindsay RC, Fisher RP, Turtle JW, and Fulero SM. 2000. From the lab to the police station. A successful application of eyewitness research. *Am Psychol* 55(6):581-598.

Dissemination of research findings

Publications

- 2017 Rawofi L, Edwards M, Krithika S, Le P, Cha D, Yang Z, Ma Y, Wang J, Su B, Jin L, Norton HL, Parra EJ. Genome-wide association study of pigmentary traits (skin and iris color) in individuals of East Asian ancestry. *Peer J* 5e2951. doi: 10.7717/peerj.3951
- 2016 Norton HL, Edwards M, Krithika S, Johnson M, Werren EA, Parra E. Quantitative assessment of skin, hair, and iris variation in a diverse sample of individuals and associated genetic variation. *American Journal of Physical Anthropology*. 160:570-581. doi: 10.1002/ajpa.22861

Presentations

- 2018 Norton HL, Fist L, Parra EJ. Associations of Fitzpatrick Skin Type with skin M index in diverse populations. *Poster presented at the American Association of Physical Anthropology meeting in Austin, TX, April 11-14, 2018.*
- 2017 Lona-Durazo F, Edwards M, Krithika S, Le P, Norton HL, Parra EJ. New regions underlying pigmentation of skin, hair and iris using quantitative phenotypes in individuals of European ancestry. *Poster presented at the Society for Molecular Biology and Evolution (SMBE) meeting 2017, Austin, TX July 2-6 2017.*
- 2017 Rawofi L, Edwards M, Krithika S, Murray N, Norton HL, Parra EJ. Genetics of pigmentation in East Asia: The role of OCA2 polymorphisms. *Poster presentation at the 86th annual meeting of the American Association of Physical Anthropologists, New Orleans, LA. April 19-22, 2017*
- 2016 Norton HL, Edwards M, Parra EJ. Quantitative characterization of pigmentation phenotype and associated genetic variation. *Paper presented at the International Union of Anthropological and Ethnological Sciences Inter-Congress. Dubrovnik, Croatia, May 4-8, 2016.*
- 2015 Norton HL, Edwards M, Werren E, Parra EJ. Skin, hair, and iris pigmentation: quantifying phenotype and identifying genetic loci associated with variation in diverse populations. *Poster presented in invited symposium ("Thinking Anthropologically About Genetics") of the 84th annual meeting of the American Association of Physical Anthropologists, St. Louis, MO. March 25-March 28, 2015.*