Shades of complexity: New perspectives on the evolution and genetic architecture of human skin

Ellen E. Quillen1,2 | Heather L. Norton3 | Esteban J. Parra4 | Frida Lona-Durazo4 | Khai C. Ang5 | Florin Mircea Illiescu6,7 | Laurel N. Pearson8 | Mark D. Shriver8 | Tina Lasisi8 | Omer Gokcumen9 | Izzy Starr9 | Yen-Lung Lin9 | Alicia R. Martin10,11,12 | Nina G. Jablonski8

1Department of Internal Medicine, Section of Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina
2Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, North Carolina
3Department of Anthropology, University of Cincinnati, Cincinnati, Ohio
4Department of Anthropology, University of Toronto – Mississauga, Mississauga, Ontario, Canada
5Department of Pathology and Jake Gittlen Laboratories for Cancer Research, Penn State College of Medicine, Hershey, Pennsylvania
6Department of Zoology, University of Cambridge, Cambridge, United Kingdom
7Centro de Estudios Interculturales e Indígenas – CIIIR, P. Universidad Católica de Chile, Santiago, Chile
8Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania
9Department of Biological Sciences, State University of New York at Buffalo, Buffalo, New York
10Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts
11Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts
12Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts

Correspondence
Ellen E. Quillen, Department of Internal Medicine, Section of Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC.
Email: equillen@wakehealth.edu

Funding information
National Science Foundation, Grant/Award Number: 1714867

Abstract
Like many highly variable human traits, more than a dozen genes are known to contribute to the full range of skin color. However, the historical bias in favor of genetic studies in European and European-derived populations has blinded us to the magnitude of pigmentation's complexity. As deliberate efforts are being made to better characterize diverse global populations and new sequencing technologies, better measurement tools, functional assessments, predictive modeling, and ancient DNA analyses become more widely accessible, we are beginning to appreciate how limited our understanding of the genetic bases of human skin color have been. Novel variants in genes not previously linked to pigmentation have been identified and evidence is mounting that there are hundreds more variants yet to be found. Even for genes that have been exhaustively characterized in European populations like MC1R, OCA2, and SLC24A5, research in previously understudied groups is leading to a new appreciation of the degree to which genetic diversity, epistatic interactions, pleiotropy, admixture, global and local adaptation, and cultural practices operate in population-specific ways to shape the genetic architecture of skin color. Furthermore, we are coming to terms with how factors like tanning response and barrier function may also have influenced selection on skin throughout human history. By examining how our knowledge of pigmentation genetics has shifted in the last decade, we can better appreciate how far we have come in understanding human diversity and the still long road ahead for understanding many complex human traits.

KEYWORDS
genetics, human evolution, skin pigmentation
1 | INTRODUCTION

Skin pigmentation is often held out among complex traits as a relatively simple phenotype to study as the melanogenesis pathway is well characterized and the phenotype comparatively straightforward to measure. This is a mischaracterization of a trait, which is, like many widely variable traits, highly polygenic with population-specific alleles, complex epistatic interactions, pleiotropic effects, and gene-specific evolutionary trajectories shaped by global and local selective pressures and drift. Much of our current knowledge of the genetic architecture of human pigmentation variation has been centered on European populations. These studies fall into two primary groups: those identifying genetic loci that explain differences in pigmentation between European and non-European populations (Bleiza et al., 2013; Bonilla et al., 2005; Lamason et al., 2005; Norton, Koki, & Friedlaender, 2007; Shriner et al., 2003) and those identifying loci that explain pigmentation variation within European populations (Candille et al., 2012; Eiberg et al., 2008; Flanagan et al., 2000; Kayser et al., 2008; Liu et al., 2015; Praetorius et al., 2013; Sulem et al., 2007; Valverde, Healy, Jackson, Rees, & Thody, 1995). While these studies built a strong foundation for our understanding of the genetics of human pigmentation variation, this Eurocentric bias has also limited our understanding of the genetic architecture of pigmentation globally. We now know that loci important in European pigmentation variation may not be relevant in non-European populations and may be small in number compared to loci influencing pigmentation outside of Europe (Crawford et al., 2017; Martin et al., 2017).

In the following pages, we will review some of the major changes in our understanding of the genetic bases of skin color as more detailed studies are carried out in more diverse populations. In doing so, we aim not to produce a comprehensive list of the genes contributing to skin color variation in humans, but to highlight the complex effects, distributions, and interactions among a few genes of particular note. Then, we will delve into how population history and selective pressures have shaped the genetic architecture of pigmentation and consider two alternative functions of skin—tanning response and barrier function—which recent evidence suggests may also exert selective pressures on skin. Finally, we will consider important technological advances in how skin color is measured and predicted that both inform and are informed by the field’s efforts to capture a more accurate global picture of skin color and genetic diversity.

2 | OLD GENES, NEW UNDERSTANDING

Much in the field has changed since the topic of pigmentation genetics was last reviewed in the Yearbook of Physical Anthropology in 2007 (Parra, 2007). At that point, the vast majority of genetic variation had only been deeply characterized in a limited number of individuals, let alone populations. The largest catalogue of common variants was the International HapMap Project which had released its data just a few years prior and focused exclusively on a small sample of white Americans with northern and western European ancestry, Yoruban individuals from Ibadan, Nigeria, Han Chinese individuals from Beijing, and Japanese individuals from Tokyo (The International HapMap Consortium, 2003). The geographic distribution of these individuals artificially created an impression of discontinuous genetic variation along (U.S.) racial lines. The Human Genome Diversity Project (HGDP) attempted to rectify this limited geographic coverage, but had less genetic data and was still hardly representative of the full geographic or genetic distribution of the human species (Cavalli-Sforza, 2005). Both projects lacked any associated phenotype data.

Over the past decade, rapidly decreasing sequencing costs alongside deliberate efforts to better explore the rich genetic and phenotypic variation within continental groups and the role that geography, history, and selection have played in shaping modern pigmentation diversity has led to a deeper understanding of the interactions among genes (Veeramah & Hammer, 2014). While the geographic distribution and ethnic diversity of populations studied has improved (Table 1), the historical bias in favor of relatively wealthy countries persists and limits our understanding of this culturally and biologically important phenotype. Figure 1 highlights the greater number of studies with large sample sizes performed in European and European-derived populations, meaning that much human genetic diversity is missed. As pigmentation genetics research is undertaken in understudied populations in Africa, South Asia, and Southeast Asia, we increasingly recognize that our understanding of the function of genetic variants is highly bound by the context in which they were first identified. The evolutionary history of the human species, which inhabits wildly diverse environments, adds an additional layer of complexity where the contribution of the same gene—or even the same allele—in determining pigmentation variation differs across populations due to interactions across many genes, local adaptation, and patterns of genetic drift and gene flow across our broadly geographically dispersed species. The comparatively limited amount of research done in non-European populations suggests that even genes like MC1R, SLC24A5, and OCA2 that have been nearly exhaustively characterized in European populations, are incompletely understood and have more to teach us about the complexity of pigmentation genetics.

2.1 | MC1R: lessons in genetic diversity

One of the most well-studied pigmentation loci in or outside of Europe is the melanocortin-1 receptor, or MC1R gene. MC1R is a seven-pass transmembrane G-protein coupled receptor found on the surface of melanocyte cells. When the α-melanocyte stimulating hormone (α-MSH) binds to the Mc1r protein, eumelanin production occurs. However, mutations that interfere with this binding process can result in the production of red-yellow phaeomelanin rather than brown or black eumelanin. Although small in size (951 bp), MC1R exhibits high levels of variation in populations across Eurasia (Harding et al., 2000; Rana et al., 1999) and pleiotropic influence on several pigmentation phenotypes. MC1R is perhaps most well-known for its role in influencing hair color, with multiple nonsynonymous variants strongly associated with red hair and fair skin color within European and European-derived populations (Box, Wyeth, O’Gorman, Martin, & Sturm, 1997; Duffy et al., 2004; Flanagan et al., 2000; Valverde et al., 1995). In addition to red hair and lighter skin pigmentation, derived alleles at MC1R are also associated with freckles, increased nevus...
<table>
<thead>
<tr>
<th>Study</th>
<th>Study type</th>
<th>Trait</th>
<th>$N$</th>
<th>Continent</th>
<th>Population</th>
<th>$N^b$</th>
<th>Associated genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abe, Tamiya, Nakamura,</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>456</td>
<td>Asia</td>
<td>Japanese</td>
<td>1</td>
<td>OCA2</td>
</tr>
<tr>
<td>Hozumi, &amp; Suzuki, 2013</td>
<td>gene (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ang et al., 2012</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>492</td>
<td>Asia</td>
<td>Peninsular Malaysia</td>
<td>2</td>
<td>SLC24A5, SLC45A2</td>
</tr>
<tr>
<td></td>
<td>gene (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beleza, Johnson, et al.,</td>
<td>Admixture</td>
<td>Melanin Index</td>
<td>699</td>
<td>Africa</td>
<td>Cape Verdean</td>
<td>4</td>
<td>SLC24A5, TYR, OCA2, SLC45A2</td>
</tr>
<tr>
<td>2013</td>
<td>mapping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candille et al., 2012</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>469</td>
<td>Europe</td>
<td>Ireland (146), Poland (72), Italy (109), Portugal (142)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Crawford et al., 2017</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>1,570</td>
<td>Africa</td>
<td>Ethiopia, Tanzania, Botswana</td>
<td>6</td>
<td>SLC24A5, MFSD12, DDB1, TMEM138, OCA2, HERC2</td>
</tr>
<tr>
<td>Eaton et al., 2015</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>419</td>
<td>Asia</td>
<td>East Asian Ancestry (Canada)</td>
<td>1</td>
<td>OCA2</td>
</tr>
<tr>
<td>Han et al., 2008</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>10,755</td>
<td>Europe</td>
<td>European Ancestry (U.S. and Australia)</td>
<td>3</td>
<td>SLC24A4, IRF4, SLC45A2</td>
</tr>
<tr>
<td>Hernandez-Pacheco et al.,</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>658</td>
<td>Americas</td>
<td>Hispanic/Latinos from Puerto Rico (285) and African Americans (373)</td>
<td>4</td>
<td>SLC24A5, SLC45A2, BEND7, PRPF18</td>
</tr>
<tr>
<td>2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonnalagadda, Norton,</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>533</td>
<td>India</td>
<td>Western India</td>
<td>1</td>
<td>SLC24A5</td>
</tr>
<tr>
<td>Ozarkar, Kulkarni, &amp; Ashma,</td>
<td>gene (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liu et al., 2015</td>
<td>GWAS</td>
<td>Self-Report</td>
<td>17,262</td>
<td>Europe</td>
<td>Dutch (5,857), Australian (4,296), UK (5,278)</td>
<td>5</td>
<td>SLC45A2, IRF4, HERC2/OCA2, MC1r, A3P</td>
</tr>
<tr>
<td>Marcheco-Teruel et al.,</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>1,019</td>
<td>Americas</td>
<td>Cuba</td>
<td>2</td>
<td>SLC24A5, SLC45A2</td>
</tr>
<tr>
<td>2014</td>
<td>gene (15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martin, Lin, et al., 2017</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>456</td>
<td>Africa</td>
<td>KhoeeSan</td>
<td>4</td>
<td>SLC24A5, TYR1, SMARCA2/ VLDL1r, SNX13</td>
</tr>
<tr>
<td>Nan et al., 2009</td>
<td>GWAS</td>
<td>Tanning response</td>
<td>15,155</td>
<td>Europe</td>
<td>European Ancestry (U.S.)</td>
<td>5</td>
<td>SLC45A2, IRF4, TYR, OCA2, MC1r</td>
</tr>
<tr>
<td>Norton et al., 2007</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>583</td>
<td>Americas</td>
<td>European Ancestry (U.S. and Canada)</td>
<td>3</td>
<td>SLC45A2, IRF4, HERC2</td>
</tr>
<tr>
<td>Norton, Werren, &amp;</td>
<td>gene (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friedlaender, 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paik et al., 2011</td>
<td>Linkage</td>
<td>Melanin Index</td>
<td>345</td>
<td>Asia</td>
<td>Mongolia</td>
<td>4</td>
<td>GRM6, AT1F1, WNT1, SILV/Pmel17</td>
</tr>
<tr>
<td>Quillen et al., 2011</td>
<td>Candidate</td>
<td>Melanin Index</td>
<td>515</td>
<td>Americas</td>
<td>Mexico (95), U.S. Hispanics (247), Colombian (173)</td>
<td>4</td>
<td>SLC24A5, SLC45A2, OPRM1, EGRF</td>
</tr>
<tr>
<td>Rawofi et al., 2017</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>305</td>
<td>Asia</td>
<td>East Asian Ancestry (Canada) &amp; Chinese</td>
<td>1</td>
<td>ZNF804B</td>
</tr>
<tr>
<td>Shriver et al., 2003</td>
<td>Admixture</td>
<td>Melanin Index</td>
<td>592</td>
<td>African/European</td>
<td>African American (232), European American (187), and African Caribbean (173 U.K)</td>
<td>2</td>
<td>TYR, OCA2</td>
</tr>
<tr>
<td></td>
<td>Mapping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Candidate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continues)
count, and an increased risk of melanoma (Bishop et al., 2009; Flanagan et al., 2000).

Two decades ago, MC1R was the only gene known to have a substantial effect on normal human pigmentation variation. The ratio of synonymous to nonsynonymous substitutions in MC1R in the genomes of humans compared to chimpanzees indicates a shared selective sweep in all modern humans which may have been associated with the loss of fur approximately 1.2 million years ago (Rogers, Iltis, & Wooding, 2004). When lightly pigmented skin under dark fur was exposed to the intense ultraviolet radiation (UVR) of equatorial Africa, a strong selective pressure would have favored darker skin in our hominin ancestors. Work on nucleotide substitutions at this locus in modern populations revealed a lack of variation in African populations and an abundance of non-synonymous substitutions in European populations. Two views were presented on the modern distribution of MC1R variation: it was argued to be under diversifying selection by some (Rana et al., 1999) and simply representing neutral drift by others (Harding et al., 2000). However, all sides agreed that the apparent lack of diversity of MC1R was evidence of functional constraint in the African continent. While functional constraint in the MC1R locus is a logical conclusion based on the evidence presented, the lack of variation was taken to mean that skin pigmentation itself was functionally constrained in Africa. This view was compounded by these studies’ reliance on African Americans and immigrants from West Africa to represent the full spectrum of the continent’s genetic and phenotypic diversity. Afro-descendant individuals have variable amounts of African ancestry that, particularly in the Americas, stems predominantly from West and Central Africa as a result of the trans-Atlantic slave trade (Bryc et al., 2010; Campbell & Tishkoff, 2010; Tishkoff et al., 2009; Zakharia et al., 2009). The use of African diaspora populations as representative of African genetic diversity is further problematic because of common adherence among scientists and study populations to the folk “one-drop rule” which ascribes Black racial identity to individuals with any amount of African ancestry regardless of multifactorial ancestry.

Over the past decade, advances in our knowledge of African skin pigmentation genetics have been driven by a shift in populations used to represent diversity on this vast continent. The earliest studies in skin pigmentation genetics used the contrast in phenotype between northern Europeans and western Africans to uncover some of the genes playing a role in human skin pigmentation. This contrast between light and dark, derived and ancestral, was useful to some extent, but limited in what it could tell us about skin pigmentation genetics within the continent of Africa, which represents the most diverse gene pool in the world—both in terms of overall genetic diversity and skin pigmentation. Although studies of admixed populations can be very useful for identifying the alleles influencing traits that vary between the parental populations, it is important to recognize that these are association studies and, as such, favor the discovery of common alleles of large effect. Given the broader variation in skin color and greater genetic diversity in Africa, relative to Europe, admixture mapping studies favor the discovery of variants that have lightened the skin of Europeans (which is explained by fewer genes) over alleles that affect skin color variation within Africa (Beleza, Johnson, et al., 2013; Bonilla et al., 2005). Nevertheless, the greater diversity in

### Table 1 (Continued)

<table>
<thead>
<tr>
<th>Study</th>
<th>Study type</th>
<th>Trait</th>
<th>Population</th>
<th>N</th>
<th>Continent</th>
<th>Associated genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stokowski et al., 2007</td>
<td>GWAS</td>
<td>Melanin Index</td>
<td>South Asian Ancestry</td>
<td>968</td>
<td>South Asian Ancestry</td>
<td>SLC24A5, SLC45A2, TYR</td>
</tr>
<tr>
<td>Sulem et al., 2007</td>
<td>GWAS</td>
<td>Skin sensitivity</td>
<td>European Iceland (7,246), Dutch</td>
<td>6,918</td>
<td>European</td>
<td>1</td>
</tr>
<tr>
<td>Visconti et al., 2018</td>
<td>GWAS</td>
<td>Tanning ability</td>
<td>European &amp; European Ancestry</td>
<td>17,6,678</td>
<td>European</td>
<td>20</td>
</tr>
<tr>
<td>Zhang et al., 2013</td>
<td>GWAS</td>
<td>Tanning ability</td>
<td>European Americans</td>
<td>9,678</td>
<td>European</td>
<td>5</td>
</tr>
</tbody>
</table>

* All genome-wide association studies (GWAS) considered more than 200,000 SNPs. The number of genes (many with more than one SNP) is indicated for candidate gene studies in parentheses.

b The number of genes significantly associated is listed. For GWAS, 10⁻⁸ is used as cutoff, for candidate gene studies, associations are listed as reported in manuscript.
African skin color and genetics is also evident in analyses of the relationship between individual admixture and skin reflectance in admixed populations as variance in skin color is greater in individuals with higher proportions of west African ancestry (Martin, Lin, et al., 2017; Parra, Kittles, & Shriver, 2004; Shriver et al., 2003).

When studies are undertaken in other regions of Africa, a more complex picture of genetic and pigmentation variation emerges. In the last few years, the focus on geographically dispersed African populations within the continent, as opposed to African-derived populations in Europe and the U.S., has identified novel genes and complex interactions contributing to African skin pigmentation. These data, though relatively new, should not be surprising given what we have long known about the deep and complex evolutionary history of African populations and the varied effects of natural selection on populations living in vastly different UVR-regimes. Even for MC1R, thought to be tightly constrained across Africa, studies of individuals from southern Africa suggest that some relatively low levels of nonsynonymous MC1R variation may be tolerated or compensated for in lower UVR regions of the continent (John, Makova, Li, Jenkins, & Ramsay, 2003).

Despite the overarching narrative of Africans as representatives of a homogeneously darkly pigmented ancestral group, some have emphasized the diversity of pigmentation across the continent (Jablonski & Chaplin, 2000; Relethford, 1997). South African KhoeSan populations, in particular, have long been highlighted as an example of convergent evolution of lighter pigmentation in Africa. However, it was not until recently that large-scale research on the genetic architecture of skin color in Southern Africa was reported (Martin, Lin, et al., 2017). These authors describe associations with loci that had been previously discovered (SLC24A5, SLC45A2, and KITLG) and a novel association at SNX13. Interestingly, they show that a large portion of the phenotypic variance in these populations remains unexplained, suggesting that skin pigmentation genomics may be far more polygenic and complex—with 50 or more genes of small to moderate effect required to explain variation within the KhoeSan—than was presumed on the basis of large-effect single nucleotide polymorphisms (SNPs), like those in MC1R, discovered in Eurasian and African European admixed populations.

Assumptions about darkly pigmented skin and functional constraints at MC1R also influenced our early understanding of pigmentation variation in populations from Southeast Asia and Oceania. Few studies have specifically focused on understanding the genetic architecture of pigmentation in these regions, so much of what we know is based on reports of allele frequencies from studies examining global variation at a particular pigmentation locus but without measured phenotypes. These studies indicate that MC1R variation in Oceanian populations may be constrained by purifying selection (Harding et al., 2000; Rana et al., 1999), while less constraint is evident in populations from Southeast Asia (Harding et al., 2000; Nakayama et al., 2006; Rana et al., 1999). Researchers also identified the presence of some intermediate frequency non-synonymous MC1R polymorphisms in both regions, including Val92Met (rs2228479, a valine to methionine substitution at amino acid 92 in the gene) and Arg163Gln (rs885479). Both derived variants are also present in Europe, but at lower frequencies. The relatively higher frequencies of the derived alleles in East Asia (23% and 64% respectively in HapMap Han Chinese) compared to Western Europe (6% and 10% in HapMap Europeans) may be due either to drift or selection following the divergence of these populations. In the case of Val92Met, this selection may have followed introgression of the variant from Neandertals (Ding et al., 2014). In the absence of pigmentation data collected in Southeast Asia and Oceania, it was impossible to know if these nonsynonymous changes were associated with pigmentation variation as assumed. Yamaguchi et al. (2012) reported a nominal association of the derived allele 163Gln with lighter skin pigmentation in a Japanese sample. However, this marker was not associated with skin pigmentation in a recent genome-wide association study (GWAS) in individuals of East Asian ancestry living in Toronto (Rawolf et al., 2017).
so further investigation is needed to confirm the effect of this polymorphism across East Asia.

The studies described above have generally included relatively limited numbers of Melanesians, making it difficult to assess if MC1R variation is constrained by purifying selection in this high-UVR region as it is in similarly high-UVR regions of Africa. This is an important question because there are many instances where recurrent exposure to either high-UVR or low-UVR climates over evolutionary history could lead to convergent depigmentation or repigmentation in human populations. Studying the genetic architecture of populations with similar levels of melanin in similar climates is important to establish the flexibility of the melanogenesis pathway. A recent investigation of melanin in similar climates is important to establish the flexibility of the melanogenesis pathway. A recent investigation of melanin in similar climates is important to establish the flexibility of the melanogenesis pathway. A recent investigation of melanin in similar climates is important to establish the flexibility of the melanogenesis pathway.

A recent investigation of MC1R sequence variation in a sample of 188 Melanesians failed to support a model of purifying selection using a McDonald-Kreitman test (Norton, Werren, & Friedlaender, 2015). This same study genotyped the Val92Met polymorphism in 635 Melanesians from 20 populations on three different islands and reported that the frequency of the derived allele ranged from 0.04 to 0.33. However, the polymorphism showed no association with skin or hair pigmentation. Taken together, these results suggest that MC1R variation, though relatively common, does not play a significant role in the genetic architecture of Melanesian pigmentation, possibly due to interactions with other unidentified loci.

An earlier study directly investigating the genetics of pigmentation variation in Island Melanesia focused on genotyping a limited set of 10 SNPs previously associated with pigmentation variation (Norton, Kittles, et al., 2007; Norton, Koki, & Friedlaender, 2007). However, these SNPs were identified primarily because they influenced variation in European populations or explained substantial differences in skin pigmentation between Europeans and other populations. It is not surprising, given this strong European ascertainment bias, that six of these polymorphisms were monomorphic or exhibited very low levels of heterozygosity across the region. Of the remaining loci (rs6058017 in the gene ASIP, rs1800404 in the gene OCA2, and Val92Met and rs2228478 in the gene MC1R), only the variants in ASIP and OCA2 showed significant associations with skin pigmentation across the region (Norton, Kittles, et al., 2007; Norton, Koki, & Friedlaender, 2007). In both cases, the ancestral allele, more common in African populations, was associated with darker skin color. However, the interpretation of these results is complicated by the strong population structure that characterizes Island Melanesia (Friedlaender et al., 2008). Both alleles exhibit significant frequency differences among islands, as does pigmentation. Without controlling for background levels of population substructure, it is difficult to determine if the observed association reflects a true influence on phenotype or is instead simply consistent with background levels of population stratification. Studies exploring the genetic architecture of skin pigmentation in Island Melanesia, or any geographically dispersed populations, should control for this structure and utilize sequencing to identify novel alleles that are not subject to a strong European-ascertainment bias.

2.2 SLC24A5: lessons on epistatic interactions

Perhaps the gene with the largest effect on European-specific decreases in skin pigmentation is solute carrier family 24 member 5 (SLC24A5). The impact of the gene on melanogenesis was identified because of its role in the lighter pigmentation that characterizes the golden mutant form of zebrafish (Danio rerio), whose melanophore-rich dark stripes are considerably lighter compared to the wild-type zebrafish (Lamason et al., 2005). Transmission electron microscopy imaging, performed to determine the cellular basis of golden hypopigmentation, showed that golden melanosomes were smaller, less electron-dense, and irregularly shaped. In humans, interruption of SLC24A5 activity similarly restricts the maturation and normal functioning of melanosomes (Ginger et al., 2008; Wei et al., 2013). A nonsynonymous mutation of an A in place of a G allele in exon 3 (Ala111Thr, rs1426654) is associated with lighter skin color in humans and explains 25-38% of variation in eumelanin levels between populations of European and African descent (Lamason et al., 2005). In addition, a striking reduction in heterozygosity in a 150-kb region near SLC24A5 in European samples indicates a history of strong selection at this locus in these populations.

Polymorphisms in a second member of the SLC family, SLC45A2 (Phe374Leu, rs16891982 and Glu272Lys, rs26722), also exhibit a high degree of differentiation between Europeans and other populations and the derived alleles of each have been associated with darker pigmentation among southern Europeans (Graf, Hodgson, & Van Daal, 2005; Norton, Kittles, et al., 2007; Norton, Koki, & Friedlaender, 2007; Soejima & Koda, 2007; Soejima, Tachida, Ishida, Sano, & Koda, 2006; Stokowski et al., 2007). The derived allele of Glu272Lys is also found at roughly 40% frequency in East Asian populations (Graf et al., 2005; Sturm, 2006; Wollstein et al., 2017). These two polymorphisms highlight the utility of functional assays in model organisms to validate associations in natural populations. As these associations were based on the distributions of genetic variants and phenotypes within and across populations, they are sensitive to the confounding effects of population stratification or linkage disequilibrium which can obscure the causative variant. Every human carries several thousand sequence variants across their genome (1000 Genomes Project Consortium, 2012; Chiang et al., 2018; Ng et al., 2008; Stenson et al., 2009; The International HapMap 3 Consortium, 2010), and the total number of known variants across all individuals and populations will only increase with the widespread adoption of low-cost sequencing technology. Elucidating the functional significance of these mutations (if any) is a major challenge to overcome given the lack of a standardized method to test the phenotypic impact of these variants. Published computational approaches do not provide a solution as they exhibit poor concordance (Tennessen et al., 2012).

Tsetskhladze et al. (2012) developed an experimental approach to assess the functional significance of specific human mutations affecting coding sequence in zebrafish and applied it first to members of the SLC gene family. Site-directed mutagenesis was used to create orthologous mutations in zebrafish cDNAs corresponding to each human variant, which were then used to generate mRNAs for rescue experiments (e.g., the two derived human variants of SLC45A2: slc45a2Phe374Leu and slc45a2Glu272Lys), using a method called
Humanized Zebrafish Orthologous Rescue (HuZOR). Zebrafish were bred with mutations knocking out the gene function and the mRNAs were injected into the mutant embryos to see if the human versions of the genes restored normal pigmentation to the albino zebrafish. The results of these studies confirmed that Phe374Leu plays an important role in skin pigmentation while Glu272Lys has a little to no effect. This suggests that the association seen by researchers may have been due to other mutations in linkage disequilibrium with the Glu272Lys polymorphism or additional genes interacting with the variant of interest. With the advancement in genome editing techniques such as CRISPR/Cas-9 technology (Blackburn, Campbell, Clark, & Ekker, 2013; D’Agostino et al., 2016; Gratz et al., 2014; Hsu, Lander, & Zhang, 2014; Hwang et al., 2013; Nissan, Perli, Fridkin, Perez-Pinera, & Lu, 2014; Seruggia, Fernández, Cantero, Pelczar, & Montoliu, 2015; Shah, Davey, Whitebirch, Miller, & Moens, 2015), functional validation will become increasingly high throughput.

The primary value of these functional validation methods is that it removes the influence of the broader genetic background in which the mutation occurs. Epistatic interactions—interactions among genetic variants at different loci which cause nonlinear effects on the phenotype—can either compound or mask the effect of a particular gene on melanogenesis. When the role of a particular gene is only understood within the confines of a single population or a group of related populations, its influence may be overstated. When the variant is studied in additional populations, the genetic variant may interact with additional variants not present in the original population that attenuate the magnitude of the original variant’s effect. For example, we have known for some time that the variants in SLC24A5 and SLC45A2 at high prevalence in Europe were also relatively common in South Asia and a GWAS in a population from the UK of South Asian descent found that both Phe374Leu in SLC45A2 and Ala111Thr in SLC24A5 are associated with skin color variation (Stokowski et al., 2007). A large scale, sequencing-based study of individuals from the southern part of India, indicates that approximately 27% of the variation in skin color within this sample is explained by Ala111Thr. Furthermore, the haplotype in India shares identity-by-descent with the variant in Europe with a coalescent time approximately 22,000–28,000 years ago (Basu Mallick et al., 2013). The impact of Ala111Thr on skin color, however, varies across the Indian subcontinent depending on the epistatic context (Jonnalagadda, Norton, et al., 2016; Mishra et al., 2017). Importantly, the rigid social structure in India characterized by high variability between populations has the potential to confound the overall effect of a variant on skin color variation when populations are combined for an association study; hence it is crucial to appropriately account for phenotypic and genetic population differences. The distribution of both pigmentation and the 111Thr allele is correlated both with geography and type of population—higher frequencies of the derived allele are seen in Northwestern India, in castes compared to tribes, and in populations with increased ancestral Northern India ancestry where haplotype analysis indicates a history of strong positive selection (Basu Mallick et al., 2013; Jonnalagadda et al., 2017). It is possible that selection occurred in North India and not in South India because of increased requirement for vitamin D synthesis or due to a strong selection event favoring lighter skin in Middle Eastern and Central Asian populations prior to population migration into India (Pagani et al., 2016).

While the SLC24A5 111Thr allele has a general skin lightening effect across Indian populations, it does not influence intra-population skin color variation in some populations such as Brahmins from Uttar Pradesh among whom the light skin allele is close to fixation or among Tibeto-Burmans where it is absent. Furthermore, the 111Thr allele shows moderate to high frequency (50%) in some highly melanized populations, the genetic variant may interact with additional variants in influence may be overstated. When the variant is studied in additional populations, the genetic variant may interact with additional variants not present in the original population that attenuate the magnitude of the original variant’s effect. For example, we have known for some time that the variants in SLC24A5 and SLC45A2 at high prevalence in Europe were also relatively common in South Asia and a GWAS in a population from the UK of South Asian descent found that both Phe374Leu in SLC45A2 and Ala111Thr in SLC24A5 are associated with skin color variation (Stokowski et al., 2007). A large scale, sequencing-based study of individuals from the southern part of India, indicates that approximately 27% of the variation in skin color within this sample is explained by Ala111Thr. Furthermore, the haplotype in India shares identity-by-descent with the variant in Europe with a coalescent time approximately 22,000–28,000 years ago (Basu Mallick et al., 2013). The impact of Ala111Thr on skin color, however, varies across the Indian subcontinent depending on the epistatic context (Jonnalagadda, Norton, et al., 2016; Mishra et al., 2017). Importantly, the rigid social structure in India characterized by high variability between populations has the potential to confound the overall effect of a variant on skin color variation when populations are combined for an association study; hence it is crucial to appropriately account for phenotypic and genetic population differences. The distribution of both pigmentation and the 111Thr allele is correlated both with geography and type of population—higher frequencies of the derived allele are seen in Northwestern India, in castes compared to tribes, and in populations with increased ancestral Northern India ancestry where haplotype analysis indicates a history of strong positive selection (Basu Mallick et al., 2013; Jonnalagadda et al., 2017). It is possible that selection occurred in North India and not in South India because of increased requirement for vitamin D synthesis or due to a strong selection event favoring lighter skin in Middle Eastern and Central Asian populations prior to population migration into India (Pagani et al., 2016).

While the SLC24A5 111Thr allele has a general skin lightening effect across Indian populations, it does not influence intra-population skin color variation in some populations such as Brahmins from Uttar Pradesh among whom the light skin allele is close to fixation or among Tibeto-Burmans where it is absent. Furthermore, the 111Thr allele shows moderate to high frequency (50%) in some highly melanized populations, the genetic variant may interact with additional variants not present in the original population that attenuate the magnitude of the original variant’s effect. For example, we have known for some time that the variants in SLC24A5 and SLC45A2 at high prevalence in Europe were also relatively common in South Asia and a GWAS in a population from the UK of South Asian descent found that both Phe374Leu in SLC45A2 and Ala111Thr in SLC24A5 are associated with skin color variation (Stokowski et al., 2007). A large scale, sequencing-based study of individuals from the southern part of India, indicates that approximately 27% of the variation in skin color within this sample is explained by Ala111Thr. Furthermore, the haplotype in India shares identity-by-descent with the variant in Europe with a coalescent time approximately 22,000–28,000 years ago (Basu Mallick et al., 2013). The impact of Ala111Thr on skin color, however, varies across the Indian subcontinent depending on the epistatic context (Jonnalagadda, Norton, et al., 2016; Mishra et al., 2017). Importantly, the rigid social structure in India characterized by high variability between populations has the potential to confound the overall effect of a variant on skin color variation when populations are combined for an association study; hence it is crucial to appropriately account for phenotypic and genetic population differences. The distribution of both pigmentation and the 111Thr allele is correlated both with geography and type of population—higher frequencies of the derived allele are seen in Northwestern India, in castes compared to tribes, and in populations with increased ancestral Northern India ancestry where haplotype analysis indicates a history of strong positive selection (Basu Mallick et al., 2013; Jonnalagadda et al., 2017). It is possible that selection occurred in North India and not in South India because of increased requirement for vitamin D synthesis or due to a strong selection event favoring lighter skin in Middle Eastern and Central Asian populations prior to population migration into India (Pagani et al., 2016).
Northern Indian populations whereby very dark skinned individuals can be found who are homozygous for what is generally considered a light skin allele, suggesting the epistatic influence of other melanin-increasing genetic variants (Iliescu et al., 2018). Studies looking at candidate genes in addition to SLC24A5 have not been able to conclusively determine a significant effect of other genetic variants on skin color variation in Indian populations (Jonnalagadda, Norton, et al., 2016; Mishra et al., 2017). What is clear however is that epistasis is present between different skin color variants as a general mechanism. In this particular instance, the SLC24A5 111Thr allele's ability to lower melanin production is dependent on specific and distinct genetic backgrounds made possible by the unique reproductive structure (caste-marriage) of the Indian populations (Iliescu et al., 2018).

Most Indian populations are the result of admixture between two main genetic ancestral components, ANI relating to West Eurasia, and ASI relating to the Andaman Islands Onge population (Auton et al., 2009; Metspalu et al., 2011; Reich, Thangaraj, Patterson, Price, & Singh, 2009). Admixture between these two ancestry components was widespread in all Indian populations up until 4,000–2,000 years ago. Following this time, cast-marriages ensued, probably due to a significant cultural shift, with different Indian populations having been reproductively isolated for at least 2000 years, up to the present (Basu, Sarkar-Roy, & Majumder, 2016; Moorjani et al., 2013). The complex landscape of skin color in India arose at the interface of natural selection, genetic drift, and social factors. West-Eurasian genetic variants at SLC24A5 and other loci interact with Indian-specific genetic variation in response to population, UVR, and social influences in the context of gene flow followed by reproductive isolation that preserved and further enhanced skin color diversity. Such highly local effects, influenced by cultural practices, are likely to have occurred in many regions of the world and represent an underappreciated evolutionary mechanism which uniquely shapes human traits.

In an attempt to understand how these modern allele distributions came to be, ancient genome studies have begun to look at the evolutionary trajectory of skin color genetic variants, such as the SLC24A5 111Thr allele, in ancient populations from Iran through Anatolia to Western and Northern Europe (Gallego-Llorente et al., 2016; Gamba et al., 2014; Hofmanová et al., 2016; Mathieson et al., 2015). These have suggested that the 111Thr was already present in Iran in the early Neolithic (10 kya) and at high frequency during the Neolithic in Anatolia, from where it arrived into Europe through population migration with the advent of agriculture 8 kya. Variants in SLC24A5 and SLC45A2 were segregating throughout much of Europe in the Mesolithic (Mittnik et al., 2018). It was thus proposed that the rapid increase in frequency of this light skin variant could be related to the advent of agriculture that led to a less vitamin D-rich diet, hence the need for lighter skin color.

Research into modern populations suggests that the 111Thr allele spread to both Southeast Asia and Northern Africa via ancient migration as well as recent admixture. In a study of three Orang Asli tribes with extremely broad distributions of skin color from peninsular Malaysia, Ang et al. (2012) identified a small but significant effect of the SLC24A5 111Thr allele associated with lighter color. However, the authors observed this allele was at low frequency (6%) and genotype at the SLC24A5 loci could explain only 1.6% of variation in pigmentation phenotype, suggesting that other, untyped loci are responsible for much of the observed variation (Ang et al., 2012).

Recent studies indicate that the SLC24A5 111Thr allele also migrated into East Africa through gene flow from the Near East (Crawford et al., 2017) and is also present among KhoeSan individuals in southern Africa at levels significantly greater than expected due to recent European admixture (Martin, Lin, et al., 2017). Long considered a highly diagnostic marker of ancestry differences nearly fixed between Northern Europeans and West Africans (often shortened in an American context to Europeans and Africans) this allele is present at frequencies up to 50% in Ethiopia and Tanzania (Pickrell et al., 2012; Schlebusch et al., 2012). These regions are geographically contiguous with the Near East and have likely shared near-constant gene flow for at least 3,000–9,000 years based on the time to most recent common ancestor (TMRCA) for the haplotypes harboring the 111Thr allele. However, the impact of SLC24A5 is not limited to northern and eastern Africa.

In both broad studies of pigmentation diversity in Africa and targeted investigations of the KhoeSan and Nama of southern Africa, the SLC24A5 Ala111Thr variant is among the most strongly associated with pigmentation (Crawford et al., 2017; Martin, Lin, et al., 2017). While the impact of this variant on skin color in these groups is clear, the effect is much smaller, explaining approximately 12% of the variation compared to the estimated 25–38% of the variation between West African and European individuals (Lamason et al., 2005). This is due in large part to the much broader distribution of skin color seen in more darkly pigmented populations and those closer to the equator as well as the greater number of loci contributing to skin pigmentation among the KhoeSan than in lightly pigmented Eurasian populations. Against the diverse genetic background of these southern African populations, the predictive models based on European populations were entirely ineffective (Martin, Lin, et al., 2017). Indeed, novel variants identified in this study of just two groups in southern Africa explained more heritable variation in pigmentation than all known pigmentation variants combined. The importance of SLC24A5 in explaining pigmentation variation has been magnified due to the relatively low amount of genetic variation in Europe and the reduced phenotypic variation in lightly pigmented populations more broadly. As a result, SLC24A5 and other genes of large effect are repeatedly identified and replicated, especially in studies of admixed populations. In the context of more genetically diverse populations such as those found in India, eastern Africa, and southern Africa, the canonical 111Thr allele, while present, explains a much smaller proportion of skin color variation while a large number of small individual effect but significant cumulative effect remain to be discovered.

### 2.3 | OCA2/HERC2: lessons in pleiotropy

Historically, skin color research was predominantly focused on the incidence and cause of albinism. In general, albinism affects ~1 in 20,000 individuals. However, rates as high as 1 in 6,500 to 1 in 28 individuals are found in some Indigenous American populations from southern Mexico, southern Brazil, eastern Panama, and the southwestern United States (Woolf, 2005). Among the Kuna residing along the Atlantic coast of Panama, albinism occurs in approximately 1 of
every 165 individuals (Keeler, 1964). This is due to a splice-site mutation in intron 17 that changes guanine to thymine in the P gene or Oculocutaneous Albinism Type 2 (OCA2), eliminating the highly conserved adenine-guanine dinucleotide 5'-splice junction, changing the reading frame, and introducing a premature stop codon in exon 19 (Carrasco, Forbes, Jeambrun, & Brilliant, 2009). A separate polymorphism, a deletion of 122.5 kb in OCA2 was found in the Navajo population causing albinism with the estimated prevalence of 1 in 1,500 to 1 in 2,000. It was estimated that the mutation originated 400–1,000 years ago from a single founder (Yi et al., 2003). In addition to the Kuna and Navajo, high rates of albinism have also been observed in Hopi, Laguna, Zuni, Jemez, and San Juan tribes, suggesting founder effects among multiple Indigenous American populations.

Such highly localized demographic effects can have an outsized influence on the distribution of both normal and disease pigmentation traits. Currently, the strongest evidence that OCA2 influences pigmentation phenotypes within European populations comes from multiple reports identifying variants in the OCA2/HERC2 complex associated with blue vs. brown eye color phenotypes (Elberg et al., 2008; Kayser et al., 2008; Sturm et al., 2008). A polymorphism in intron 86 of HERC2 (rs12913832) functions as an enhancer element influencing OCA2 expression in human melanocytes (Visser, Kayser, & Palstra, 2012). The ancestral T allele at this locus is associated with increased OCA2 expression and brown eyes, while the derived C allele leads to reduced OCA2 expression and blue eyes. The vast majority of unrelated, blue-eyed individuals from throughout Europe, Turkey, and Jordan carry a single haplotype containing the C allele, suggesting a common origin of this mutation. Initial reports suggested that a founding mutation causing blue eyes likely occurred in the Near East and spread throughout Europe with the migration of agriculturalists out of this region during the Neolithic. The recent availability of genome sequences from hundreds of ancient Europeans has allowed unprecedented resolution of this region's genetic history and challenged those early conclusions. Studies of genomes from remains dating up to 8,000 years ago reveal that most present-day Europeans derive ancestry from west European hunter-gatherers, among whom the derived OCA2/HERC2 allele was already at high frequency, as well as early farmers from Anatolia and northern Eurasians, among whom the derived allele frequency varied between zero and 50% (Lazaridis et al., 2014; Mathieson et al., 2015; Mittnik et al., 2018). Over the past 2,000 years, the ancestors of modern Britons were under further strong selection at OCA2/HERC2, among other loci, favoring blue eyes and blond hair (Field et al., 2016).

Variation of iris and hair pigmentation outside of Europe has rarely been studied, partly due to the absence of accurate methods to quantify variation on these traits, which may seem quite homogenous in non-European populations. Norton and colleagues identified significant mean differences in hair color between individuals residing in the U.S. and Canada who self-identify as East Asian, African American, European, Hispanic, or South Asian (Norton et al., 2015). Subsequent analyses in the same East Asian individuals confirmed the presence of substantial iris color variation, some of which is explained by OCA2 polymorphisms distinct from those found in Europe (Edwards et al., 2016; Rawofì et al., 2017). One of these markers, rs1800414 (a histidine to arginine substitution at position 615, His615Arg), also has a strong effect on skin pigmentation. This pleiotropy has not frequently been documented but is expected given the shared melanogenesis pathway between eye, hair, and skin pigmentation.

The recent investigations of OCA2 support a role for the His615Arg substitution and a second independent polymorphism in the same gene (rs74653330, Ala481Thr) in the lighter skin color seen among Chinese, Japanese, and other East Asian populations (Abe et al., 2013; Eaton et al., 2015; Edwards et al., 2010; Murray, Norton, & Parra, 2015). The frequency of 615Arg is greater than 50% in East Asians while the ancestral allele is fixed in Africa and Europe and OCA2 shows clear signatures of selection in East Asia (Chen, Hey, & Slatkin, 2015; Donnelly et al., 2011; Yuasa et al., 2007). Edwards and colleagues estimated that each copy of the 615Arg allele decreases pigmentation by 0.85–1.3 melanin units in this sample of East Asian individuals. Functional experiments in human foreskin tissue and mice indicated that the derived homozygous genotype (GG) of His615Arg decreases melanin production, and functional studies in zebrafish also supported the important effect of this locus in pigmentation (Yang et al., 2016). Moreover, estimates based on several different approaches place the appearance of the derived allele of His615Arg after the split of Europeans and Asians (Chen et al., 2015; Smith, Coop, Stephens, & Novembre, 2018; Yang et al., 2016), reflecting the independent evolution of light skin pigmentation in between these populations. While OCA2 clearly impacts skin, iris, and hair color in East Asia, it explains only a very small proportion of the total variation across this vast continent. Much larger and more detailed studies are imperative to reveal the likely complex genetic architecture similar to that seen in Europe and just beginning to be revealed in Africa.

2.4 | MFSD12: a lesson in novel associations

We are still in the early days of research into African skin pigmentation genomics, but two recent studies have investigated the genetic architecture contributing to skin color variation within diverse African populations. Crawford et al. (2017) identified TMEM138, DDB1, and MFSD12 as important genes in skin pigmentation among a diverse set of predominantly eastern and southern African groups. Martin, Lin, et al. (2017) determined that SMARCA2/VLDLR and SNX13 are associated with variation within southern Africans. While some of these genes have known functions within the melanogenesis pathway, none had previously been linked to variation within or among human populations. Given the vast genetic and pigmentation diversity within Africa, that such variants exist is not surprising. In fact, Martin, Lin, et al. (2017) estimate that less than 20% of the genetically determined (heritable) variation in skin color among the KhoeSan is explained by the top 50 previously associated skin pigmentation genes.

The influence of MFSD12 on pigmentation and the geographic distribution of its alleles is particularly interesting because it upends many commonly held assumptions about the evolution of skin color. MFSD12 is conserved throughout vertebrates (Madej, Dang, Yan, & Kaback, 2013); and low mRNA levels are associated with depigmented skin in vitiligo patients (Yu et al., 2012). A heterozygous mutant mfsd12a zebrafish showed reduced protein presence in xanthophores, the cells equivalent to mammalian melanocytes. In mouse melanocytes, knockdown of Mfsd12 mRNA by ~80% via shRNAs caused a
30–50% increase in melanin content and higher density of melanosomes. The mechanism of action for this, however, appears to be through indirect action in the membranes of melanocyte lysosomes rather than melanosomes.

The derived MFSD12 alleles at rs56203814, a synonymous mutation, and intronic rs10424065 are associated with dark pigmentation and present only in African populations (Crawford et al., 2017). The derived alleles are most common in East African populations with Nilo-Saharan ancestry while the ancestral alleles are associated with light skin pigmentation and are nearly fixed in East Asians and Europeans. In addition, the San, Ethiopian, and Tanzanian populations with Afro Asiatic ancestry also carry the ancestral alleles. This finding runs strikingly counter to the widely held assumption that ancestral alleles at pigmentation-associated loci would all be associated with darker pigmentation. In the region of the genome containing TMEM138 and DDB1, variants associated with both lighter and darker skin color have been segregating for up to 600,000 years—well before the evolution of Homo sapiens. The derived, dark pigmentation-associated allele at rs7948623 in TMEM138 and the derived, light pigmentation-associated allele at rs11230664 in DDB1 have both undergone directional selection in Africans and non-Africans. These results by Crawford and colleagues highlight that there are many ways to make enough melanin for a particular UVR regime. Furthermore, phenotypic evidence suggests perhaps a threshold necessity can be met for skin pigmentation and that other factors (sexual selection, drift, etc.) can drive skin to become more darkly pigmented, not just more lightly pigmented.

3 | NATURAL SELECTION AND HUMAN CHOICE

Our profiles of MC1R, OCA2/HERC2, SLC24A5, and MFSD12 have touched on the roles of population history and local adaption in shaping allele frequencies, creating novel epistatic interactions, and regulating pigmentation in sometimes counterintuitive ways. However, it is useful to think more broadly about the evolutionary history of skin color and forces of nature and culture that shape the global distribution of skin color today.

3.1 | Directional selection and convergent evolution

Pigmentation genes are frequently over-represented in genome-wide searches for evidence of selection, revealing the particularly strong directional selection acting on this phenotype (Hancock et al., 2011; Pickrell et al., 2009; Sabeti et al., 2007; Voight, Kudaravalli, Wen, & Pritchard, 2006; Williamson et al., 2007; Wollstein & Stephan, 2015). While early evidence of selection on pigmentation comes from the functional constraints on MC1R variation in West Africa (Harding et al., 2000; Rana et al., 1999), the majority of these studies have focused on Europe. Differences in allele frequencies at Ala111Thr in SLC24A5 between the European and other HapMap populations were some of the largest observed in the genome (Lamason et al., 2005) and evidence for selection favoring variants associated with lighter skin color has been observed for other loci as well, including TYR, TYRP1, KITLG, and OCA2/HERC2 (Beleza et al., 2012; Field et al., 2016; Lao, de Grujter, van Duijn, Navarro, & Kayser, 2007; Norton, Kittles et al., 2007; Norton, Koki, & Friedlaender, 2007; Sabeti et al., 2007). The overall picture that has emerged is one in which lighter skin evolved in Europe relatively recently, with the onset of selection favoring alleles in the loci above sometime in the last 20 kya, driven by strong selective pressure (presumably to maximize vitamin D synthesis potential).

The growing number of ancient genomes available from Mesolithic, Neolithic, and Bronze age sites in Europe and Western Asia has also contributed to the discussion about the timing of the shift to lighter skin color in European populations, with several studies using genomic information from pigmentation loci to reconstruct the phenotypes of sampled individuals (Gonzalez-Fortes et al., 2017; Günther et al., 2018; Mathieson et al., 2015; Mittnik et al., 2018; Olalde et al., 2014). Using data from only 2–3 loci (SLC24A5, SLC45A2, and OCA2/HERC2) darker skin color was inferred for Mesolithic individuals from Spain, Scandinavia, and central Europe (González-Fortes et al., 2017; Günther et al., 2018; Olalde et al., 2014). In contrast, Brace et al. (2018) used a much larger set of 36 SNPs and a newly developed algorithm (Walsh et al., 2017) to predict the pigmentation phenotype of Mesolithic and Neolithic individuals from Britain. Their analyses suggest that the skin color of both individuals was likely dark, with that of Mesolithic Cheddar Man predicted to be “dark or dark to black”. These findings suggest that lighter skin color was uncommon across much of Europe during the Mesolithic. This is not, however, in conflict with the date estimates of <20 kya above, which addresses the onset of selection and not time of fixation of favored alleles (Beleza et al., 2013; Beleza, Johnson, et al., 2013). While ancient genome studies predict generally darker skin color among Mesolithic Europeans, derived alleles at rs1426654 and rs16891982 were segregating in European populations during the Mesolithic (Gonzalez-Fortes et al., 2017; Günther et al., 2018; Mittnik et al., 2018), suggesting that phenotypic variation due to these loci was likely present by this time. However, reconstructions of Mesolithic and Neolithic pigmentation phenotype using loci common in modern populations should be interpreted with some caution, as it is possible that other as yet unexamined loci may have also influenced phenotype.

Quantitative assessment of skin pigmentation shows that East Asians, like Europeans, tend to fall at the lighter end of the continuous global pigmentation distribution (Norton, Edwards, et al., 2015), and individuals residing at northern latitudes on both continents receive insufficient UVR to produce previtamin D$_{3}$ for at least 1 month per year (Jablonski & Chaplin, 2000, 2017; Chaplin & Jablonski, 2009). There is substantial and incompletely overlapping skin pigmentation variation within each of these populations, yet genetic contributions to variation in skin color within East Asian populations has been investigated substantially less than in European populations. The evolution of depigmentation in East Asians and the genetic mechanisms involved were first investigated by identifying putative signatures of selection in Europe and East Asia in publicly available polymorphism data (Hider et al., 2013; Izagirre, García, Junquera, De La Rúa, & Alonso, 2006; Lao et al., 2007; McEvoy, Beleza, & Shriver, 2006; Myles, Somel, Tang, Kelso, & Stoneking, 2007; Norton, Koki, & Friedlaender, 2007; Norton, Koki, & Friedlaender, 2007; Norton, Kittles, et al., 2007). The parsimonious explanation would be that these
low-UV dwelling populations shared common selective sweeps and some evidence for this was found in KITLG and OCA2. However, the vast majority of selection happened either specifically on European haplotypes (SLC24A5, SLC45A2, and TYR) or on East Asian haplotypes (DCT, ADAM17, ADAMT20, OCA2, BNC2, TYPR1, LYST, and MC1R). This example of convergent adaptation to low-UVR environments is in contrast to the variants in MC1R which are shared by both darkly pigmented Melanesians and Sub-Saharan Africans residing in high-UVR environments. While only a few studies of pigmentation have been done in East Asian populations, it is clear the segregating polymorphisms associated with lighter skin pigmentation in East Asia are distinct from those in Europe.

While there is strong evidence for selection acting to decrease pigmentation levels in both European and East Asian populations, this has not resulted in a homogeneous pigmentation phenotype in either region. Some indigenous populations of northern Europe and Asia living at higher latitudes (and so experiencing lower UVR levels) have darker pigmentation than populations living further to the south, where UV is higher and the facility for pre-vitamin D3 synthesis should be greater (Jablonski & Chaplin, 2000). This non-intuitive pattern is also observed in indigenous populations of the Americas, where it has been explained as a relaxation of selection for lighter skin color due to the significant amount of foods high in vitamin D3 present in traditional diets (Jablonski & Chaplin, 2010). The darker pigmentation observed in these European and Asian populations may be attributable to a number of factors that may include plentiful dietary sources of pre-vitamin D3 as well as demographic processes (e.g., low admixture rates with populations further to the south) that would have limited the spread of alleles such as Ala111Thr. To date, there has been comparatively little work exploring skin pigmentation phenotypes and associated genetic variation in these northern European and Asian populations, leaving this an area for open investigation.

A long-standing question in the evolution of human skin color is whether the darker pigmentation common in populations living in high UVR regions, including Africa, South Asia, and Melanesia, is due to shared ancestral alleles or if darker pigmentation, like lighter pigmentation, has evolved multiple times in recent human history. Crawford et al. (2017) reported that the haplotypes at loci associated with darker pigmentation in several regions of Africa are shared with southern Asian and Australo-Melanesian populations, supporting a common-origin model for the evolution of darker skin pigmentations in these populations. The authors also observed that both light and dark alleles at MFSD12, DDB1, OCA2, and HERC2 emerged prior to the origin of modern humans and roughly equal proportions of light and dark alleles are ancestral, suggesting that both types of alleles may have been segregating in early human populations migrating throughout Africa and the rest of the world. This raises the possibility that these loci may explain some of the variation in pigmentation phenotypes observed in equatorial climates outside of Africa.

Populations from Island Melanesia, including the countries of Papua New Guinea and the Solomon Islands, generally exhibit darkly pigmented skin consistent with expectations given the high UVR levels across the region. However, in a study of 1,135 individuals from several islands across the Bismarck Archipelago of Papua New Guinea, Norton and colleagues reported that pigmentation is highly variable (with Melanin Index values ranging from 50 to 115), with much of that variation structured along geographic lines (Norton et al., 2006). Using phenotypic data collected from a wide range of populations, Martin and colleagues demonstrated that this pattern, that is, generally darker pigmentation but with a wide variance in pigmentation levels, is characteristic of equatorial populations (Martin, Lin, et al., 2017). They argue that this pattern is consistent with the “threshold model” (Chaplin, 2004), in which natural selection favors darker skin color, but above a certain minimum value pigmentation phenotype varies with little to no fitness cost. This “threshold model” has interesting ramifications for expectations of diversity in and around pigmentation loci. Under a model of strong purifying selection nonsynonymous polymorphisms leading to a decrease in pigmentation are expected to be rare in equatorial populations. However, under the threshold model such alleles could be tolerated as long as other loci ensured sufficiently high levels of melanin. This may partly explain the observed patterns of MC1R sequence diversity described in Island Melanesia and other incomplete selective sweeps at pigmentation loci.

### 3.2 Demographic factors and mating practices

While continent-scale selective sweeps have been relatively well-characterized, the more subtle interplay between natural selection and the uniquely human impact of culture and demography in shaping pigmentation diversity within smaller regions remains poorly understood. Indeed, some studies suggest that classical positive selection may have a relatively limited effect in recent human evolution (Hernandez et al., 2011). Instead, more complex and variable evolutionary phenomena may explain the observed genetic variation among humans (Key, Teixeira, de Filippo, & André, 2014), including introgression of adaptive alleles, balancing selection, local positive selection, reduced purifying selection, and selection on regions involving structural variants, among others. The impact of local adaptation and culturally limited mating practices is evident in India where skin color varies from very light skin, similar to observations in European populations (with measures of MI < 30) to very dark skin, comparable to observations in populations from sub-equatorial Africa (MI > 80; Basu Mallick et al., 2013; Jonnalagadda, Ozarkar, Ashma, & Kulkarni, 2016; Mishra et al., 2017). While at the global level UVR explains more than 80% of skin color variation (Chaplin, 2004), there is no clear correlation between geography and skin color within India and UVR is responsible for only 16% of Indian skin color diversity (Iliescu et al., 2018). This suggests that natural selection for skin color was weaker in India and mitigated by numerous population admixture and migration events. Throughout India, both lighter and darker skin populations can be found in all regions with social status a stronger predictor of skin color than geography such that the population to which an individual belongs could explain 42% of skin color variation within the country (Iliescu et al., 2018). Preference for within caste marriage over numerous generations has reinforced these genetic differences and led to neighboring populations exhibiting significantly different skin color profiles, with Brahmin populations showing generally lighter skin color.

Researchers have also investigated the role of sexual selection in the evolution of human skin color, particularly to account for variation within similar UVR climates (Aoki, 2002; Robins, 1991; van den...
Berghe & Frost, 1986). Global patterns of sexual dimorphism—a possible outcome of sexual selective pressure—suggest that the impact of sexual selection on pigmentation is weak (Madrigal & Kelly, 2007). However, it may have targeted local effects in places like Island Melanesia and India, where sexual dimorphism in skin color is pronounced (Iliescu et al., 2018; Norton et al., 2006). Although global patterns of sexual selection have not been found, they are not necessarily to be expected as it would require uniform preferences. However, the role of culturally determined mating practices as well as sexual selection can have potent local effects modifying the influence of UVR-determined directional selection.

3.3 | Admixture: genes in new contexts

Considerable variation in skin pigmentation is observed among Indigenous American populations reflecting the complex history of multiple shifts in UVR environment over the lineage. As humans migrated into the New World, within the past 20,000 years, selection for skin darkening was likely among groups that moved into areas of more intense UVR near the tropics (Jablonski & Chaplin, 2000, 2010; Relethford, 1997). This geographic patterning is diffused in the current populations of the Americas by generations of recent admixture among geographically distant populations resulting in broad variation in skin color and complex underlying genetic architecture. Each country in the Americas has a unique history of admixture including indigenous peoples, European colonists, and, in many countries, African slaves and 19th and 20th century migrants from East Asia and South Asia. The genetic diversity across the Americas reflects the idiosyncratic history of mating among these groups resulting in a broad range of interactions among genetic variants related to pigmentation. As a result, considerable population structure is seen both among and within Latin American countries (Homburger et al., 2015; Moreno-Estrada et al., 2014; Ruiz-Linares et al., 2014).

Compared to the substantial body of research focused on differences in skin pigmentation between Europe and Africa, relatively few studies have investigated populations of the Americas despite evidence from admixed populations that there is a positive correlation between Indigenous American ancestry and darker pigmentation. To date, published studies of genetic contributions to skin pigmentation variation in the Americas have focused on candidate gene approaches and ancestry associations in admixed Latino and Latin American populations. The first reported pigmentation associations for Hispanics and Latinos were in ancestry informative loci in the MYOS, CYP19A1, and SLC24A5 genes (Hoggart et al., 2003). Quillen and colleagues employed a selection-nominated candidate gene approach in Mexican Americans and Colombians with admixed European and Indigenous American ancestry. They reported associations in well supported pigmentation candidates, SLC24A5 and SLC45A2, and novel associations in OPRM1 and EGFR in these admixed populations (Quillen et al., 2011). Recent research suggests the influence of OPRM1 on pigmentation may be mediated by sun-seeking behavior (Khouja, Lewis, & Bonilla, 2018).

In their 2014 publication, the CANDELA (Consortium for the Analysis of the Diversity of Evolution of Latin America) project tested the association of polymorphisms in nine pigmentation candidate genes for two admixed Brazilian study population, the Gaúcho and Baiano. The authors report significant associations between SNPs in pigmentation candidates SLC24A5, SLC45A2, HERC2, and TYR and Melanin Index when the two admixed groups were analyzed as a combined dataset. It is important to note that the study populations are composed of a three-way admixture between European, African, and Indigenous American ancestries. Both the Gaúcho and Baiano have similarly low contributions of Indigenous American ancestry, approximately 8%. Additionally, only a weak correlation was found between Melanin Index and Indigenous American ancestry, suggesting that the genotype and pigmentation associations reported may be largely driven by the European/African component of admixture (De Cerqueira et al., 2014). In three-way models of admixture in the Americas, the substantial genetic and phenotypic variation contributed by the (western and southern) African component frequently masks the influence of Indigenous American ancestry on skin color (Bonilla, Shriver, Parra, Jones, & Fernández, 2004). While well-characterized panels of ancestry informative markers (AIMs) for Indigenous Americans have been published that are useful for admixture mapping analyses (Galanter et al., 2012), to date, only limited panels of fewer than 150 AIMs have been used in admixture association studies in Indigenous American populations (De Cerqueira et al., 2014; Klimentidis, Miller, & Shriver, 2009; Quillen et al., 2011; Ruiz-Linares et al., 2014).

Compared to North and South America, even less research has focused on the impact of both recent and pre-Colombian migration on patterns of pigmentation in the Caribbean. Torres, Stone, and Kittles (2013) examined 420 individuals from eight islands using 105 autosomal AIMs to estimate proportional biogeographic ancestry. On average, these African Caribbean populations exhibit 77% West African, 15% Europeans and 7.7% Indigenous American ancestry but the variation within and across these geographically compact islands is substantial (Torres et al., 2013). Populations from St. Kitts has the lowest Indigenous American ancestry (5.8%), lowest European ancestry (8.2%), and highest African ancestry (85.9%) among the populations studied. Less than 260 km away, the Kalinago from Dominica are reported to have the highest Indigenous American ancestry among Caribbean islanders at 61% with 31% West African and 8% European ancestries (Ang et al., 2017). Despite complications inherent in modeling three-way admixture, some studies of skin pigmentation have successfully been undertaken. Hernandez-Pacheco et al. (2017) performed a GWAS among Puerto Rican individuals which identified 82 loci suggestively associated with skin color. However, because there have been relatively few large-scale studies of Hispanic or Latino populations where skin color was quantitatively measured, replication had to be performed in an African American sample with a strikingly different admixture profile. The results replicated only three major regions including SLC24A5 and SLC45A2 and a novel intergenic region between BEND7 and PRPF18.

In the largest study of skin color genetics in the Caribbean to date, 1,019 individuals sampled from all the provinces of Cuba were genotyped using 128 autosomal AIMs to estimate admixture (Marchecro-Teruel et al., 2014). This population has 72% European, 20% West African and 8% Indigenous American ancestries on average with significant variation across the island. The highest Indigenous American ancestry was reported at 15% from the province of Granma...
and highest reported African ancestry was observed in the provinces of Guantánamo (40%) and Santiago de Cuba (39%). Sixteen previously documented skin pigmentation SNPs were analyzed while controlling for local genetic ancestry and four of them were significantly associated with Melanin Index: rs1426654 (SLC24A5 Ala111Thr), rs16891982 (SLC45A2 Phe374Leu), rs35395 (SLC45A2), and rs12913832 (HERC2/OCA2). All of these are known European-ascertained variants. This candidate SNP approach is a first step in understanding the role of these variants in these populations; however, admixed individuals are of particular interest because descendants of populations with fixed differences in allele frequencies may exhibit polygenotypes that have not previously been documented. Without genome-wide studies, we will remain in the dark as to how these unique interactions may be shaping skin color.

4 | DYNAMIC RESPONSES AND MOVING BEYOND SKIN COLOR

In our focus thus far on the genetic bases of skin color, we have relied, as the vast majority of researchers do, on an unspoken definition of skin color as the genetically determined amount of melanin produced by the body in the absence of UV exposure (constitutive skin color). This is most commonly measured on the underside of the proximal arm, an area rarely exposed to the sun. The underlying assumption of this focus is that constitutive pigmentation is the primary object of selection both as a primary mediator of fitness and the aspect of skin most directly under genetic control. However, there are a number of additional selective pressures that may have influenced skin’s evolutionary history to a greater or lesser degree, including metabolic conservation (Elias & Williams, 2016), arid climates (Elias, Menon, Wetzel, & Williams, 2010), and temperature (Yang et al., 2018). We will focus in greater detail on two: the tanning capacity of skin and its role as a barrier.

4.1 | Tanning as the primary mediator of UVR exposure

Skin pigmentation is the primary mediating factor in preventing UVR-induced damage in humans, but it is not static. Facultative pigmentation, the amount of melanin in sun-exposed skin, is a more accurate measure of the ability of the body to protect itself. Facultative pigmentation is more challenging to measure, however, because it is determined by genetic as well as short-term and long-term environmental exposures. Nevertheless, lab-based studies with controlled UVR exposure have identified ancestry-associated differences in the dose–response upregulation of melanogenesis. Wagner and colleagues demonstrated clearly distinct patterns of tanning intensity between European American and Latino individuals, suggesting ancestry-associated differences in tanning response persist after constitutive pigmentation is incorporated into the analysis (Wagner, Parra, Norton, Jovel, & Shriver, 2002). Additional indirect evidence as to the genetic architecture of constitutive and facultative pigmentation comes from measures of heritability of these phenotypes. Among KhoeSan individuals, the heritability of facultative pigmentation—measured as the difference between exposed and unexposed skin—is roughly one-third that of constitutive pigmentation highlighting that both genetic variation and a significant environmental component contribute to this trait (Martin, Lin, et al., 2017).

The literature on genetic regulation of tanning response, even more so than for constitutive pigmentation, is dominated by studies of European populations. This focus is due, in part, to the increased risk of skin cancer among these populations following repeated DNA damage following UVR exposure. This is reinforced by the findings from genome-wide association analyses which have predominantly identified genetic variants already known to influence constitutive pigmentation in Europeans. However, meta-analyses of genes contributing to pigmentation, sun sensitivity, and/or skin cancer risk highlight the substantial degree to which these do not overlap (Gerstenblith, Shi, & Landi, 2010).

As in studies of constitutive pigmentation, MC1R was among the first genes identified as contributing to tanning response with the same variants linked to lighter skin color, red hair, and freckles also linked to a decreased tanning capacity. Recent research extends these findings to show sex-by-genotype interactions on tanning that are likely mediated by the fact that MC1R encodes a receptor for estrogen-dependent α-melanocyte stimulating hormone (Hernando et al., 2016). While lighter pigmentation does increase risk of DNA damage, this association with melanoma is due not just to MC1R’s impact on constitutive pigmentation phenotype, but also to its independent role in influencing the DNA repair process (Kadekaro et al., 2005, 2010; Robbins et al., 1993). Many other genes implicated in GWAS in European populations similarly show pleiotropic effects. Interestingly, many of these sun-sensitivity variants were detected alongside hair or eye color variation, likely because these vary to a greater degree within European populations than skin color. Derived variants in IRF4, SLC24A4, and HERC2/OCA2 were linked to lighter skin color, less tanning ability, and blue/light eye color in a study of British and Australian individuals of European ancestry (Han et al., 2008). Similarly, many known variants associated with hair color were replicated and found to be associated with sun sensitivity among Europeans and European Americans including ASIP, SLC45A2, IRF4, TYR, OCA2, and MC1R (Nan et al., 2009; Sulem et al., 2008).

In the largest study to date, a GWAS was performed on 176,000 individuals of European ancestry based on self-reported tanning ability (Visconti et al., 2018). This study confirmed associations in HERC2/OCA2, IRF4, MC1R, RALY/ASIP, SLC45A2, and TYR as well as pigmentation genes not previously associated with tanning: BNC2, TPCN2, SLC24A4, TYP1, and DCT (TYRP2). Additional genes not previously linked to skin color but known to regulate MITF were identified, in addition to several genes of unknown role in melanogenesis. While a link to melanogenesis has not been documented, at least one of these genes plays a role early in melanocyte development. These results indicate that variants contributing to facultative melanogenesis may exhibit their influence during development, broadly across regulatory signaling pathways, or in the well-established melanogenesis pathway. Very few studies have examined facultative pigmentation in admixed or non-European populations. A linkage- and association-based study of quantitatively assessed facultative tanning in a Mongolian population implicated regions of the genome containing GRM6, ATF1, WNT1, and...
While these are all known pigmentation candidate genes, it is noteworthy that none of them were associated with tanning in Europeans. Additional studies including East Asian individuals are essential for a more complete understanding of the genetic architecture of tanning as long-lasting facultative pigmentation following UVR exposure is more common among individuals of southeast Asian ancestry (Choe, Rim, & Youn, 2002; Jo, Yoon, Woo, & Youn, 2006). A possible mechanism for this was identified in a study of tanning ability in Korean women which reported a previously unknown pigmentation gene, WW domain-containing oxidoreductase (WWOX), is associated with tanning ability. Silencing of this gene in cell culture dramatically increases tyrosinase activity, reflecting its potential functional importance in melanogenesis.

Virtually nothing is known about tanning among lightly or moderately pigmented African or American populations. A recent summary of the existing literature on population-level differences in tanning response advanced the hypothesis that highly labile facultative pigmentation among lightly to moderately pigmented individuals could form the basis for a convergent adaptation to UVB exposure at latitudes where seasonal variation is substantial (Quillen, 2015). While UVA causes acute DNA damage and triggers the release of stored melanin, UVB is the primary driver of increased melanogenesis over days and weeks. This may be particularly important as repeated exposure to narrow-band UVB radiation significantly depletes serum folate whereas exposure to UVA does not (Gambichler, Bader, Sauermann, Altmeyer, & Hoffmann, 2001; Shaheen, Abdel Fattah, & El-Borhamy, 2006). The strength of this selective pressure would depend heavily on the balance between photoprotection and photodamage, an unresolved question (Solano, 2016).

### 4.2 The barrier function of skin

While skin color and its function in regulating UVR is widely discussed among anthropologists, characteristics of skin such as structural integrity, thickness, developmental trajectories, and permeability against pathogens and allergens, have been studied primarily within a medical context, without much appreciation for their evolutionary bases and anthropological implications (Bergboer, Zeeuwen, & Schalkwijk, 2012; Boguniewicz & Leung, 2011). Recent work has highlighted genes related to skin barrier function evolving rapidly and under adaptive forces in the human lineage (Goodwin & de Guzman Strong, 2016). With the availability of large genomic and other "omics" databases, it is now possible to investigate the barrier function of skin, especially its role in innate immune function and as a host to the human microbiome within an anthropological context.

One of the challenges of studying skin-barrier function is the sheer complexity of this system: thousands of genes are expressed at varying levels in the different layers of skin, and have a variety of structures, functions, and interactions (Edqvist et al., 2015). On top of this "host" complexity, skin is a dynamic ecosystem of bacteria, viruses, and fungi dependent on each other and the skin environment (Byrd, Belkaid, & Segre, 2018). One of the most anthropologically relevant questions is: how does human genetic, cultural, and environmental diversity shape this complex system? Rather than a complete answer to this question, we catch glimpses, mainly from the medical literature. For example, major trends in the skin microbiome depend heavily on broader environmental conditions rather than host genetic makeup (Byrd et al., 2018). Nevertheless, host genetic variation along with microbiome composition creates strong predispositions toward immune-mediated skin disorders, such as psoriasis and atopic dermatitis (Chng et al., 2016; de Cid et al., 2009; Lenz et al., 2015; Palmer et al., 2006; Yan et al., 2017). These studies have begun to elucidate the underlying foundations of skin disease etiologies, but the natural variations of skin-barrier function and their adaptive roles remain mostly unknown.

One major contribution of epidemiological studies in regard to skin barrier function is to pinpoint a single region of the genome referred to as the epidermal differentiation complex (EDC). This ~2 Mb region on chromosome 1 harbors dozens of genes from multiple gene families, all of which primarily function in the development and activity of human skin (Strasser et al., 2014). As such, the EDC is a prime candidate for anthropological studies to understand inherited skin barrier-function variation among humans. For example, RPTN in the EDC was highlighted as one of the very few "human-accelerated" genes as compared to the Neandertal genome (Green et al., 2010).

One of the important features of the EDC is its structural complexity. Based on evolutionary analyses of structural variants across the genome, Lin and colleagues found that the EDC is among the regions with the highest density of structural variants (Lin & Gokcen, 2018; Lin, Pavlidis, Karakoc, Ajay, & Gokcen, 2015), some which appear to be adaptively evolving. A previous genome-wide study found that a large deletion overlapping two EDC genes, LCE3B and LCE3C, predates Human-Neandertal divergence (Lin et al., 2015). This 32 kb-long deletion is a strong susceptibility risk factor for psoriasis (de Cid et al., 2009), yet remains common (~50%) among human populations. Considering its age, it is unclear why purifying selection has not eliminated this variant. In contrast, the haplotype block harboring the deletion retains high allele frequency among extant and ancient human populations, contains an excess of variants with intermediate frequency, and has an unusually long history compared to the rest of the genome. These results are best explained by the LCE3B/C deletion being maintained under balancing selection in humans. A biological hypothesis for this balance is that the affected healing process resulting from this deletion variant predisposes a carrier to skin disease, but at the same time boosts the person's immunity (Pajic, Lin, Xu, & Gokcen, 2016).

Additional structural variants like filaggrin (FLG), one of the largest genes in the genome with copy number variable exonic repeats, serves as a natural candidate as it directly contributes to skin barrier function and is a major risk factor of atopic dermatitis (Palmer et al., 2006; Sandilands, Sutherland, Irvine, & McLean, 2009). Some findings suggest that FLG loss of function is adaptively correlated with latitude as a response to changes in UVR exposure (Thyssen, Bikle, & Elias, 2014). More recent work, however, suggests that FLG is neutrally evolving in the human lineage, potentially on its way for pseudogenization (Eaaswarkhanth et al., 2016). This same work found an unusually long haplotype block (Huxian haplogroup) overlapping FLG, that appears to have undergone a selective sweep in East Asians. This seeming contradiction (neutrality on FLG but a selected haplotype in the same locus) is likely explained by the Huxian haplogroup
regulating neighboring EDC gene, Hornerin (HRNR). While the function of HRNR is not clear, the Huxian haplogroup is significantly correlated with the microbiome composition of healthy skin, implying that the HRNR might alter the skin microbiome. The question why HRNR function is under selection in East Asian populations remains an ongoing area of inquiry.

The EDC is a region where complex and diverse evolutionary forces, rather than classical adaptive models, may explain the variation observed in this locus. Anthropological insights into how human skin has evolved are key to the formulation of hypotheses and interpretation of population genetics analyses. For example, the timing of different human technologies, pathogenic pressures and exposure to different environmental factors will provide clues as to why variations like the LCE3BC deletion and Huxian haplotype were favored across human evolution. For both skin color and barrier function, anthropologically contextualized samples from populations exposed to different environmental conditions (e.g., variable humidity, UVR, pathogenic pressures) provide an important framework for comparative studies of genetic variation.

5 | TECHNOLOGY AND NEW FRONTIERS

The rapid growth in pigmentation genetics, or much of genomics, could not have been possible without high-throughput, relatively low-cost genotyping and sequencing methods which allow for the discovery of new variants in more populations. Such technological change occurs so rapidly that it frequently dictates how studies are performed. In the pursuit of larger sample sizes and higher resolution genotyping, too often the quality of phenotypes collecting is ignored and samples of convenience are used, limiting the generalizability of findings.

5.1 | Measuring melanin

In this new era of genomic studies of pigmentation, the importance of precise methods for evaluating phenotypes is consistently undervalued. Given the many genes of small effect contributing to complex traits, it is imperative that pigmentation be measured accurately. There have been dramatic advances in the technologies used to evaluate pigmentation in human populations. For skin pigmentation, one of the first innovations was the introduction of reflectance spectrophotometers in the 1950s, such as the EEL instrument widely used by European researchers (Barnicot, 1958; Harrison & Owen, 1964; Weiner, 1951), or the Photovolt primarily employed by Americans (Garn, Selby, & Crawford, 1956; Lasker, 1954). Both portable instruments provided quantitative measures of skin pigmentation based on spectrophotometry, which is much more reliable than previous methods based on subjective visual matching, such as the colored glass tiles developed by Felix von Luschan that were widely used in the first half of the 20th century (Thomas, 1905; von Luschan & von Luschan, 1914). While these quantitative measures were a significant advancement, measures from the machines were not directly equivalent, necessitating the development of a series of formulas to convert between the systems (Garrard, Harrison, & Owen, 1967; Lees & Byard, 1978; Lees, Byard, & Relethford, 1979). Efforts have similarly been made to develop methods for converting measures taken on the von Luschan tiles into equivalent reflectometry quantitative measures (Swiatonowski, Quillen, Shriver, & Jablonski, 2013).

Further improvements in terms of portability have taken place in the last decade. Handheld portable instruments, such as Cortex Technology DSM Colormeter series (DSM II and DSM III) or Konika Minolta spectrophotometers (CM2500D, CM2600D, CM-600D, CM700D) or Chromameters (CR-300, CR-400, CR-410) have been widely used in recent anthropological and dermatological studies. In recent years, quantitative measures of pigmentation have been typically reported using either the CIE 1979 L*a*b* (CIELAB) color space (CIE, 1986; Weatherall & Coombs, 1992) or the Melanin and Erythema Index (Dawson et al., 1980; Diffee, Oliver, & Farr, 1984). CIELAB is an international standardized color system designed to approximate human color vision. In this system, color is reported in three dimensions: a brightness dimension (L*), a green/red dimension (a*) and a blue/yellow dimension (b*). The Melanin Index was developed to specifically measure melanin content in the skin, and it is based on reflectance in the red spectrum, where hemoglobin, another major chromophore of the skin, does not absorb light. The Erythema Index is a measure of the degree of redness of the skin due to hemoglobin and is calculated by subtracting the absorbance in the red filter from the absorbance of the green filter. A high correlation has been reported in studies that compared L* and Melanin Index, and a* and Erythema Index, in the same individuals (Shriver & Parra, 2000; Wagner, Jovel, Norton, Parra, & Shriver, 2002; Yun et al., 2010). However, there is evidence indicating that the L* color dimension may be to some extent influenced by the redness of the skin (Fullerton et al., 1996; Shriver & Parra, 2000; Takiwaki, Overgaard, & Serup, 1994). Furthermore, there is little reverse compatibility between CIE measures, M/E Index, and the earlier EEL and Photovolt machines.

Although quantitative descriptions of skin pigmentation have primarily been reported as L* and Melanin Index values based on readings from reflectometers, there are many other alternatives. For example, it is possible to describe the optical properties of the skin using diffuse reflectance spectroscopy (DRS), which measures reflectance at specific intervals along the entire visible spectrum (400–700 nm). This approach can assess individual skin chromophores (e.g., melanin, oxy- and deoxy-hemoglobin), and provides a more nuanced picture of pigmentation than colorimetric methods (Seo, Kim, Kim, Kye, & Ahn, 2011), but has not been widely adopted in the anthropological field. Another promising approach is the quantification of skin pigmentation based on digital photographs. In a recent study, Walsh and colleagues classified individuals from four different samples in three or five skin color categories based on the evaluation of digital photographs by an experienced dermatologist (Walsh et al., 2017). It is also possible to extract the red, green, blue (RGB) and luminosity measurements from digital photographs to obtain quantitative measures of pigmentation in the CIELAB color space (Coelho, Miller, Zmudzka, & Beer, 2006; Seo et al., 2011) or as Melanin Index (Majewski et al., 2016). A critical aspect to consider in the application of digital photography is the standardization of all the relevant parameters, such as camera, camera settings, distance and lighting conditions. A color standard, such as the widely used Macbeth ColorChecker Card (McCamy, Marcus, & Davidson, 1976) can be placed next to the skin to identify and correct for any potential effects of variable illumination (Anderson, Hallam, Nduka, & Osorio, 2015).
improvements in the accuracy and ease with which skin color measurements can be obtained are essential in advancing the breadth of populations in which skin color can be studied, a fundamental barrier to understanding the global genetic architecture of skin color. However, it is important for researchers to be aware of the unique limitations of each system and the ways that the data can and cannot be compared to previously collected data.

5.2 Predicting skin color

The many advances over the last decade in our understanding of genetic variants contributing to pigmentation variability within and between diverse human populations have enabled the development of predictive models that are often motivated by forensic science. These models have evolved over time and vary considerably in approach and accuracy (Table 2). These models primarily use probabilistic classifiers that consider individual-level genotypes associated in prior studies to predict pigmentation as a discrete trait. Rather than evaluating the accuracy of these models in predicting Melanin Index or another continuous value, most work to date has evaluated prediction accuracy in categories, such as “Very Pale, Pale, Intermediate, Dark, and Dark-Black” skin color (Walsh et al., 2017). However, assessing prediction accuracy with continuous phenotypes is straightforward and more interpretable using quantitative reflectance measures (Liu et al., 2015; Valenzuela et al., 2010), and dichotomization of quantitative measures is rarely defensible and often misleading. Thus, future work would benefit from prediction and accuracy assessment from models of clearly continuous phenotypes, especially for skin pigmentation, the most genetically complex and continuous of the pigmentation phenotypes (MacCallum, Zhang, Preacher, & Rucker, 2002).

Unlike for other complex phenotypes in human genetics, polygenic scores have not gained traction; this is for good reasons, as a relatively small number of variants can explain a significant fraction of the variation in skin pigmentation for many populations (Beleza, Johnson, et al., 2013). Furthermore, there are limitations in applying polygenic scores across diverse populations due to sample ascertainment and population structure, especially for an adaptive trait (Berg & Coop, 2014; Martin et al., 2017; Scutari, Mackay, & Balding, 2016).

A striking feature of pigmentation genetics relative to other complex traits in humans is the clear role that epistasis plays in determining trait variation. This has implications both for predictive models as well as for assessing heritability in a population. Specifically, modeling epistatic effects can improve genetic prediction accuracy (Alipanahi, Fontanillas, 23andMe Research Team, Pitts, & Gentleman, 2017; Pośpiech et al., 2014). Additionally, because the most widely used SNP-based heritability software, GCTA, is based on purely additive effects (Bulik-Sullivan et al., 2015; Yang et al., 2010; Yang, Lee, Goddard, & Visscher, 2011), unconsidered epistatic effects are expected to inflate heritability estimates (Young & Durbin, 2014). The effects of some specific epistatic interactions of large effect have been functionally validated. For example, an intronic variant in HERC2 (rs12913832) modulates eye color by functioning as a distal enhancer that regulates OCA2 transcription via chromatin-loop formation (Visser et al., 2012). Additionally, a common variant (rs12203592) in IRF4, a transcription factor with no known role in melanocyte biology, is associated with pale skin, eyes, and brown hair; this variant alters transcription factor binding of MITF (melanocyte master regulator), which in turn modulates pigmentation by altering TYR expression (Praetorius et al., 2013). Next generation sequencing will play an important role in fine-mapping coding and regulatory variants as a precursor to functional follow up of putatively causal variants, especially in trans-ethnic analyses (Kichaev & Pasaniuc, 2015). Pinpointing these causal variants and their interactions will aid our understanding of the regulatory network influencing pigmentation variation. This, in turn, will improve both prediction and our understanding of the evolution of this globally varying phenotype.

6 SKIN: THE QUINTESSENTIAL HUMAN TRAIT

Biological anthropologists have been fascinated with skin color since the beginning of the discipline and humans have been fascinated by it much longer. This wonder is well-placed as the magnitude of skin color variation and the speed with which pigmentation-related variants have swept to high frequencies in new environments is truly remarkable. As a biomedically and forensically relevant trait, a complete picture of the genetic architecture of pigmentation requires sequencing and high-quality phenotyping of many individuals from geographically diverse populations. For individualized prediction to be effective—be it for skin cancer risk, vitamin D deficiency, psoriasis, or forensic reconstructions—it is insufficient to understand the role that a given allele plays in a single population. The early focus on genes like MC1R, which varied significantly within Europe but was constrained by purification selection elsewhere, misled us into thinking that other genes would follow that pattern. Through deliberate, but still insufficient study of diverse populations around the world, we have come to realize that many alleles will be private to a population or region and each genetic variant will have a unique distribution shaped by the demographic and cultural history of the populations in which it arises. Effect size estimates and phenotypic predictions have repeatedly failed to generalize to global populations, because the influence of the same variant will be shaped by its epistatic interactions with other variants within any given individual or population. Perhaps most importantly, as a culturally relevant trait, understanding the complex genetic architecture of pigmentation serves as a way of understanding how the human species came to be the way it is today. Through ancient and modern DNA analysis, functional studies, and consideration of the multiple evolutionary pressures exerted on skin, we can reconstruct our history. The key lessons we have learned from studying the genetic architecture of skin color over the past decade stretch beyond this not-so-simple phenotype and should inform all our studies of complex traits varying among human populations.

ACKNOWLEDGMENTS

This work is an outgrowth of the symposium “Here comes the sun: evolutionary responses to solar exposure” organized by EEQ and NGJ at the 86th annual meeting of the American Association of Physical Anthropologists. OG was funded by NSF Award Number 1714867.


climate-mediated selective pressures in humans.
Human Genetics

genetic associations in stratified populations.
Nature Academy of Sciences

1186/1471-2148-13-150

amongst direct descendants from Neolithic Aegeans.
Journal of Dermatological Science

Differences in tanning ability and sensitivity to sunlight between sexes.
American Journal of Human Genetics, 82, 36–43. http://doi.org/10.1016/j.ajhg.2007.10.003


44548. http://doi.org/10.1038/srep44548


