

# Genetic Admixture, Self-Reported Ethnicity, Self-Estimated Admixture, and Skin Pigmentation Among Hispanics and Native Americans

Yann C. Klimentidis,<sup>1\*</sup> Geoffrey F. Miller,<sup>2</sup> and Mark D. Shriver<sup>3</sup>

<sup>1</sup>*Department of Anthropology, University of New Mexico, Albuquerque, NM*

<sup>2</sup>*Department of Psychology, University of New Mexico, Albuquerque, NM*

<sup>3</sup>*Department of Anthropology, Pennsylvania State University, University Park, PA*

**KEY WORDS** ethnic identity; New Mexico; genetic ancestry

**ABSTRACT** The relationship between ethnicity and biology is of interest to anthropologists, biomedical scientists, and historians in understanding how human groups are constructed. Ethnic self-identification in recently admixed groups such as Hispanics, African Americans, and Native Americans (NA) is likely to be complex due to the heterogeneity in individual admixture proportions and social environments within these groups. This study examines the relationships between self-identified ethnicity, self-estimated admixture proportions, skin pigmentation, and genetic marker estimated admixture proportions. These measures were assessed using questionnaires, skin color measurements, and genotyping of a panel of 76 ancestry informative markers, among 170 Hispanics and NAs from New Mexico, a state known for

its complex history of interactions between people of NA and European (EU) ancestry. Results reveal that NAs underestimate their degree of EU admixture, and that Hispanics underestimate their degree of NA admixture. Within Hispanics, genetic-marker estimated admixture is better predicted by forehead skin pigmentation than by self-estimated admixture. We also find that Hispanic individuals self-identified as “half-White, half Hispanic” and “Spanish” have lower levels of NA admixture than those self-identified as “Mexican” and “Mexican American.” Such results highlight the interplay between culture and biology in how individuals identify and view themselves, and have implications for how ethnicity and disease risk are assessed in a medical setting. *Am J Phys Anthropol* 000:000–000, 2008. ©2008 Wiley-Liss, Inc.

Most human genetic variation is found within socially defined groups (Lewontin, 1972), and human diversity is better described by clines than it is by sharp boundaries (Jorde and Wooding, 2004; Barbujani, 2005). However, it is possible to make some inferences about group membership, such as in the case of recent admixture that occurs between groups that shared a common ancestor many thousands of years ago (Shriver et al., 1997). The resident population of the United States is composed of many individuals who have mixed geographic ancestry, such as African and Hispanic Americans. While individuals in these admixed groups might self-identify into a single ethnic category, highly varied admixture proportions are often found (Pfaff et al., 2001). For example, self-identified African-Americans show anywhere between 0 and 80% European ancestry (Parra et al., 1998, 2001; Sinha et al., 2006), and self-identified Hispanics show anywhere between 0 and 90% Native American (NA) admixture (Bertoni et al., 2003; Bonilla et al., 2004a; Gonzalez Burchard et al., 2005). In fact, the term “Hispanic” refers to a nationally, geographically, and ancestrally heterogeneous group produced by 400 years of biological and cultural admixture between Europeans, NAs, and West Africans (Bertoni et al., 2003; Gonzalez Burchard et al., 2005). The US census acknowledges that Hispanics occupy an ambiguous position in the American “racial” landscape, since they are assumed to have important subdivisions with respect to culture and ancestry (US Census, 2000). Given this heterogeneity, Hispanics provide an excellent opportunity to disentangle the sociocultural and biological factors that interact

to shape social identities, phenotypes, and genotypes (Gonzalez Burchard et al., 2005; Choudhry et al., 2007).

The American Southwest is especially well suited to genetic admixture analysis, since the admixture is essentially dihybrid, between people of Spanish and Southwest US Native American ancestry. These correspond closely to the parental populations that are used as part of our genetic admixture analysis. Genetic contributions from other parts of the world have been relatively low (Nieto-Phillips, 2004). For example, the West African admixture component in a neighboring sample of Hispanics living in southern Colorado was only 3.2% (Bonilla et al., 2004a).

Additional Supporting Information may be found in the online version of this article.

Grant sponsors: RPT, SRAC, GRD (Intramural Grants, University of New Mexico).

\*Correspondence to: Yann C. Klimentidis, Department of Anthropology, University of New Mexico, Albuquerque, NM, 87130 USA. E-mail: yann@unm.edu

Received 9 February 2008; accepted 1 September 2008

DOI 10.1002/ajpa.20945  
Published online in Wiley InterScience  
(www.interscience.wiley.com).

New Mexico, in particular, represents a unique opportunity to examine the sociocultural and biological determinants of Hispanic self-identification, since there is a strong cultural distinction between self-identified “Spanish” and “Mexican” groups, and a strong sense of “blood purity” in the former (Gutierrez, 1991; Nieto-Phillips, 2004). Many self-identified Spanish New Mexicans consider themselves direct descendants of the Spanish colonists who arrived in the mid to late 1500s and encountered Pueblo, Navajo, and Apache tribes. The ethnic label “Spanish American” is however a relatively recent term in New Mexico, stemming from a surge in ethnic consciousness in the mid 1800s. It served partly as a way of countering the American characterization of New Mexicans as a mixed-blood people. This newfound identity was born of the desire of many residents to demonstrate a nonmixed European heritage during New Mexico’s bid for statehood that lasted 60 years, until 1912. Currently, this ethnic label serves to distinguish New Mexican Spanish Americans from recent Mexican immigrants in terms of geographical origin, language, physical appearance, and “blood” (Gonzales, 1997; Nieto-Phillips, 2004).

One goal of this study is to determine whether detailed self-identified ethnic categories and self-estimated levels of admixture proportions correspond to marker-based genetic admixture proportions across individuals. To date, the research on these topics has been limited. Williams et al. (1992) compared genealogically self-estimated admixture and genetically measured admixture among a sample of Gila River NAs in Arizona. The measure of self-estimated admixture that they use, however, relied on the subject’s description of his/her pedigree, followed by researcher adjustments to account for the average expected European admixture of a Mexican-American, a European-American or an African-American ancestor. This method yielded a strong relationship between their measures of genetic admixture and the adjusted pedigree-estimated genetic ancestry. Although Williams et al. cross-validated these measures, their method relied on a limited number of genetic markers and an indirect measure of self-estimated admixture proportions. More recently, Bonilla et al. (2004a) found that among Hispanics in the San Luis Valley in southern Colorado, self-identified Mexican-Americans showed more NA admixture than self-identified Spanish-Americans. Sweeney et al. (2007) showed that genetic admixture corresponded to self-reported admixture among Hispanics of the Southwest. This study focused, however, on combinations of broad “race” categories such as White, Hispanic, and American Indian, and did not investigate the relationship within self-designated Hispanic subgroups.

Ethnic identification by self and others could be mediated by physical appearance. One potential phenotypic cue to assessing ethnic identity is skin color, which has historically been used for this purpose in the US. Because of this historical usage, skin color no doubt continues to shape peoples’ perceptions of ancestry. Indeed, skin color is also a reliable indicator of genetic admixture proportions in some ethnic groups. Among African-Americans and Hispanics, constitutive skin pigmentation is positively correlated with African and NA genetic admixture, respectively (Shriver et al., 2003; Bonilla et al., 2004a,b; Parra et al., 2004). However, a study among Brazilians found that a collection of physical traits, including skin pigmentation, as well as hair and

facial features, was a weak predictor of African genetic ancestry, as determined by a panel of 10 AIMs (Parra et al., 2003). Another study in Brazil also highlighted this tenuous relationship (Marrero et al., 2005). In this study, individuals who were classified as White had significant nonwhite genetic ancestry. Taken together, these results demonstrate a complex relationship between genetic and phenotypic measures of ethnicity, and suggest that the process of self-identification may be similarly confounded.

In this study, we hypothesize that 1) New Mexican Hispanics who use different sub-ethnic self-identification labels will have corresponding differences in genetic admixture, 2) New Mexican Hispanics and NAs are able to accurately estimate their own genetic admixture proportions, and 3) there is a positive relationship between NA admixture and skin pigmentation. We use a questionnaire, skin color measurements, and ancestry informative markers (AIMs) to test these hypotheses.

## SUBJECTS AND METHODS

### Study population

We recruited 170 Hispanics ( $n = 147$ ), NAs ( $n = 15$ ), and mixed ethnicity ( $n = 8$ ) out of ~700 students from introductory Psychology courses at the University of New Mexico in Albuquerque. Subjects were recruited through an online registration system, whereby students could receive class credit for participation in research studies. There were 59 males and 111 females, ranging in ages from 18 to 22. The study recruitment message stated that only Hispanics and NAs could participate, though some subsequently identified as a mixture of one of these and White. Of the self-identified Hispanics in this sample, 78% were born in New Mexico, and 95% completed at least one-third of their primary schooling in New Mexico. Among self-identified NAs, 87% were born in New Mexico, and 100% completed at least one-third of their primary schooling in New Mexico. All participants gave written informed consent, and the study was approved by the University of New Mexico Human Research Review Committee.

### Ancestry informative markers

Cheek swabs were collected from 185 participants. DNA was extracted from the swabs and purified using established protocols, and a panel of 76 biallelic AIMs was typed (see Supporting Information). The AIMs are described elsewhere (Bonilla et al., 2004a,b; Choudhry et al., 2006; Martinez-Marignac et al., 2007), and were chosen because they exhibit large allele frequency differences (denoted “ $\delta$ ”, where  $\delta = |\text{freq of “A” allele in population 1} - \text{frequency of “A” allele in population 2}|$ ) between three major continental population groups: Europeans, West Africans, and NAs (see next section for more detail). The average population pairwise  $\delta$ s for these markers are 0.47 between Europeans and West Africans, 0.44 between Europeans and NAs, and 0.53 between NAs and West Africans. Of 185 initially recruited subjects, 14 were eliminated from further analysis, because 50 or fewer AIMs were successfully genotyped (likely due to sample quality/purity), and one was eliminated for self-identifying as “White,” leaving a final sample size of 170 subjects.

TABLE 1. Average individual and group admixture percentage by self-identified ethnicity ( $\pm$ SD for the MLE individual admixture analyses) for the trihybrid and dihybrid analyses

	n	Trihybrid analysis			Dihybrid analysis	LEADMIX (dihybrid)
		Native American admixture (%)	European admixture (%)	African admixture (%)	Native American admixture (%)	Native American admixture (%)
Hispanics	147	32.7 $\pm$ 12.4	61.6 $\pm$ 12.4	5.7 $\pm$ 6.0	35.5 $\pm$ 12.1	35.0
Native Americans	15	71.8 $\pm$ 18.6	25.3 $\pm$ 19.0	2.9 $\pm$ 3.6	72.5 $\pm$ 18.9	72.8

### Parental populations

To assess the genetic ancestry from each of the three parental populations, genotype frequencies were obtained from 1) 72 individuals of Spanish origin, from the province of Valencia, Spain; 2) 177 individuals from the NA Cheyenne, Pima, Pueblo, and Maya; 3) 279 individuals from the Central African Republic, Nigeria, and Sierra Leone. Continental average allele frequencies from these samples for the 76 AIMs are presented in Supporting Information. We assume that the parental allele frequencies have not changed significantly because of drift, gene flow, or selection over the last few hundred years (Long, 1991; Wang, 2003). We use a method of estimating group admixture that accounts for genetic drift in the parental and hybrid populations. However, other recent factors, such as disease and a resulting reduction in population size, especially among NAs, may have changed these allele frequencies, such that the frequencies observed today do not accurately reflect those of 400 years ago, at the time of initial admixture. Finally, gene flow within the Americas and within Europe (i.e., from other groups in the region) may have affected the allele frequencies over the past 400 years. However, in the case of NAs, studies have shown that the allele frequencies of the AIMs like those used in this study do not tend to differ substantially across current-day populations of the Americas (Bonilla et al., 2004a; Luizon et al., 2008).

### Genetic admixture estimates

Individual genetic admixture levels were calculated using the maximum likelihood estimation (MLE) approach described by Hanis et al. (1986). Given ancestral allele frequencies at a locus, the probability of observing a marker genotype is computed for each locus. The logs of the individual locus probabilities at all loci are then summed. For every possible admixture proportion from 0 to 100, the probability of the observed genotype is computed. The admixture proportion that corresponds to the maximum combined probability across all loci is the one that is the MLE of ancestry for that individual (Halder, 2005). This program also calculates 95% confidence intervals for each estimate. We perform a trihybrid (African, European, and NA) as well as a dihybrid (European and NA) analysis. Other statistical methods used to estimate individual admixture (e.g., STRUCTURE and ADMIXMAP) typically show high degrees of correlation with the MLE method (Bonilla et al., 2004a; Martinez-Marignac et al., 2007). To estimate group admixture levels for Hispanics, NAs, and subethnic groups of Hispanics, we also used a maximum likelihood method as implemented in LEADMIX (Wang, 2003). This method accounts for genetic differentiation in parental populations prior to the admixture event, as well as genetic drift and sampling error in the parental popu-

lations and the hybrid populations subsequent to admixture.

### Measures of self-identification

Participants were asked to check a box on a questionnaire indicating whether they consider their ethnicity to be White, Hispanic, NA, or Other. Then, participants were asked: "What would you say your ethnicity is (if you could describe it any way you wanted)?" This open-ended question was followed by a long response line that allowed subjects to write anything they wanted. With this open-ended format, as opposed to a force-choice format, we hoped to capture all of the nuances in how individuals identify themselves. Indeed, we obtained a wide variety of responses to this question, and were only able to classify some of them by a few major categories. Finally, participants were asked: "On a scale of 0 to 100 with 100 being pure Native American by blood/genetics, and 0 being no Native American ancestry by blood or genetics, what would your estimate be as to your degree of Native American ancestry?" They were asked to circle 1 of 21 possible values representing increments of 5% from 0% NA/100% European to 100% NA/0% European.

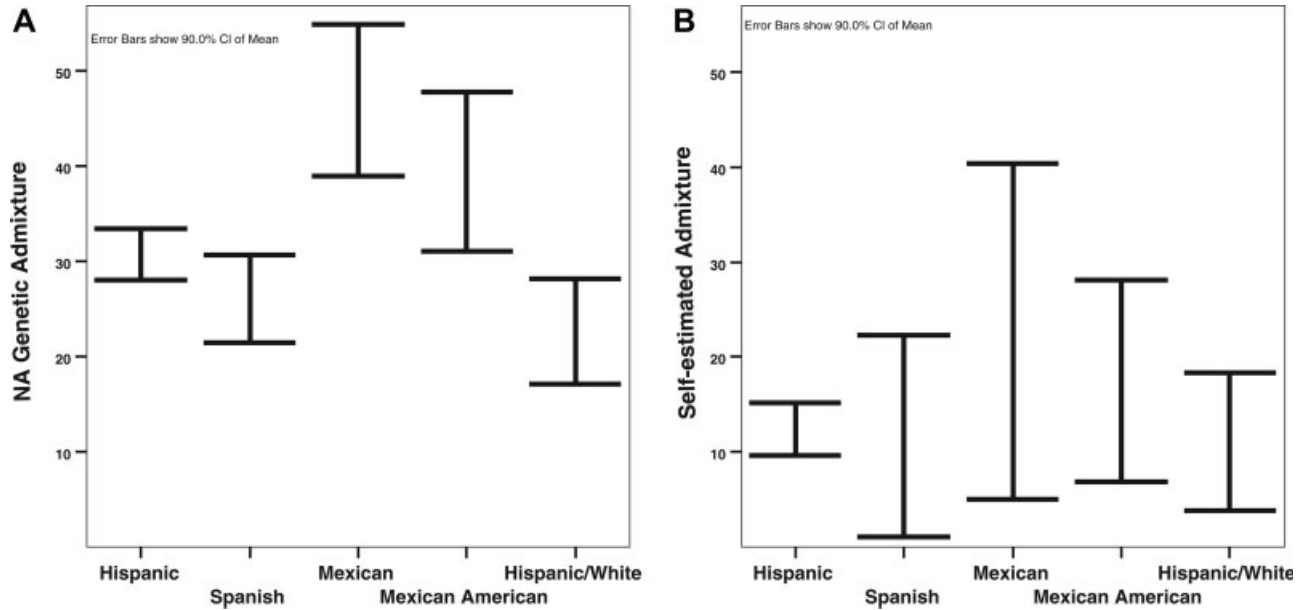
### Skin pigmentation

Skin pigmentation was measured on the upper inner arm and on the forehead using a computerized handheld narrow-band reflectometer called the DermaSpectrometer (Cortex Technology; Hadsund, Denmark). The use of forehead reflectance has been used in several recent anthropological studies (Shriver and Parra, 2000; Brutsaert et al., 2004), and can provide some assessment of tanning status when compared with less exposed body areas like the inner arm. Additionally, facultative pigmentation as measured on the forehead might vary with genetic admixture and environment differently than upper inner arm pigmentation does, since tanning capacity and/or exposure to sun also varies with genetic admixture.

For each subject, three successive measurements were taken in each location and then averaged. Because of technical problems with the reflectometer, measurements were obtained for only 104 of the 170 subjects.

### Statistical analyses

Differences in average NA and African genetic admixture between self-identified Hispanics and NAs were tested using Mann-Whitney *U* tests. To test for differences in admixture between sub-ethnic classification groups, Mann-Whitney *U* tests were performed between the NA genetic and self-estimated admixture for those who wrote in "Hispanic" and for those who wrote in each of the other classifications ("Spanish", "Mexican", etc.). To determine the relationship between NA genetic admixture and self-estimated admixture, Spearman cor-



**Fig. 1.** (a) Comparison of average Native American genetic admixture by sub-ethnic category among Hispanics; (b) Comparison of average Native American self-estimated admixture by sub-ethnic category among Hispanics.

relation coefficients were calculated for the entire sample, and then separately for the self-identified Hispanics and self-identified NAs. To examine the relationship between skin pigmentation and NA genetic admixture, Spearman correlation coefficients were calculated for the entire sample, and then separately for self-identified Hispanics and self-identified NAs.

## RESULTS

### Individual admixture estimates

Admixture estimates are examined separately for self-identified Hispanic and NA individuals. Among self-identified Hispanics, the average NA admixture is  $32.7\% \pm 12.4\%$  (see Table 1), slightly lower than the 34.1% found by Bonilla et al. (2004a) in their southern Colorado sample. Among self-identified NAs, the average NA admixture is  $71.8\% \pm 18.6\%$ , a value significantly larger than observed in Hispanics ( $P < 0.001$ ). As expected, West African admixture is low for both groups:  $5.7\% \pm 6.0\%$  for Hispanics and  $2.9\% \pm 3.6\%$  for NAs, and not significantly different between them ( $P = 0.074$ ). Both dihybrid analyses yielded highly similar results to the trihybrid analysis ( $r^2 = 0.95$ , see Table 1). We find that the 95% confidence intervals of the MLEs include values that are  $\sim \pm 13$  percentage units away from the estimate. This means that, for example, NA admixture greater than or equal to 13% is required to conclude with 95% confidence that there really is NA admixture as opposed to there being none. For the LEADMIX analyses, which take into account population differentiation and genetic drift, the 95% confidence intervals are  $\sim \pm 5\%$ .

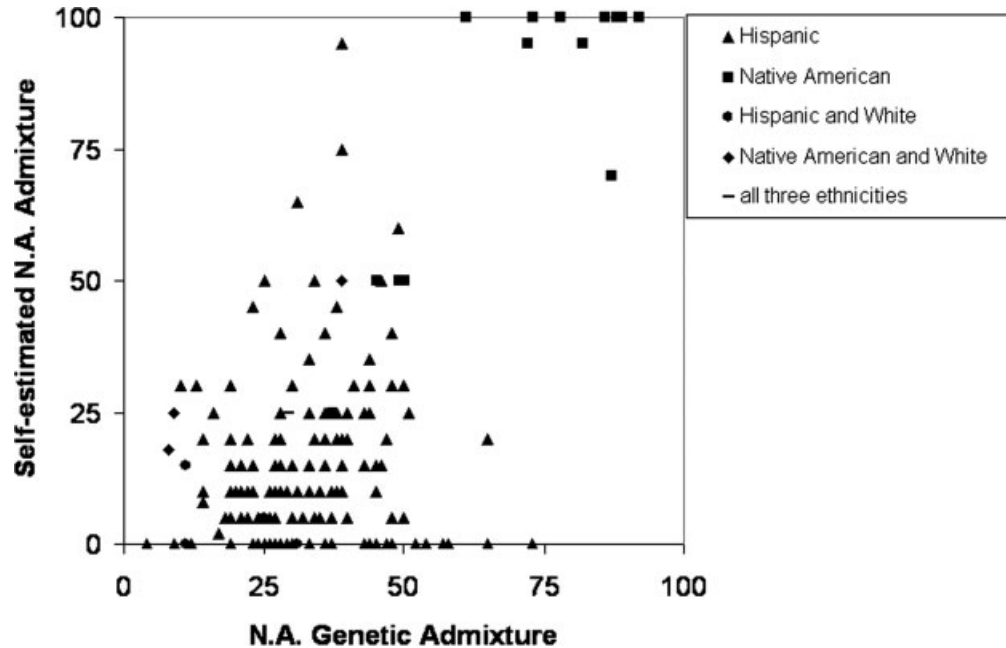
### Ethnic subclassification

Apart from making a forced-choice decision to identify themselves as White, Hispanic, NA, or Other, participants also wrote in an ethnic self-description on a long response line (in response to the question: "What would you say your ethnicity is (if you could describe it any

way you wanted)?"). Among the 147 self-identified Hispanics, approximately two-thirds of them used one of the following five self-description phrases: "Hispanic", "Spanish", "Mexican", "Mexican American", and "half Hispanic/half White". We compared both the genetic and self-estimates of NA admixture between these groups. There are larger differences between groups based on their mean individual degree of genetic admixture (Fig. 1a) than on their mean self-estimated degree of NA admixture (Fig. 1b). The Mann-Whitney test was significant for all pairwise comparisons of mean NA genetic admixture proportions (Fig. 1a), except for the Hispanic-Spanish, Mexican-Mexican American, and Spanish-Hispanic/White comparisons. There were no significant pairwise differences between any of the sub-ethnic categories for the self-estimated admixture proportions (Fig. 1b). The same holds when using the dihybrid admixture analysis. The LEADMIX program also yielded similar results for NA genetic admixture: (Hispanic: 30.6%, Spanish: 26.1%, Mexican: 49.8%, Mexican American: 39.0%, Half Hispanic/Half White: 23.1%). Thus, although there are significant differences in genetic ancestry measured using AIMs across the ethnic self-descriptions, these are not reflected in individuals' self-estimated admixture levels.

### Self-estimated admixture vs. genetic admixture

Figure 2 shows the relationship between NA genetic admixture (as determined by the 76 AIMs) and self-estimated NA admixture, with different symbols representing different self-identified ethnicities. The plot demonstrates that many individuals underestimate their respective ethnic admixture proportion. Among Hispanics,  $\sim 84\%$  (86%, using the dihybrid analysis) of individuals underestimate the amount of NA admixture estimated from the AIMs. These individuals are located under the diagonal line expected in a perfect positive correlation plot in Figure 2. Among only self-identified NAs, the correlation coefficient ( $\rho = 0.727$ ;  $P = 0.002$ ) between self-



**Fig. 2.** Self-estimated Native American admixture as a function of genetically determined Native American admixture using the trihybrid model in Hispanics and Native Americans Spearman correlation  $\rho = 0.30$ ,  $P < 0.001$ ; for the dihybrid model: Spearman correlation  $\rho = 0.29$ ,  $P < 0.001$ .

estimated admixture and genetic admixture is much higher than it is among self-identified Hispanics ( $\rho = 0.106$ ;  $P = 0.209$ ). Using the dihybrid admixture model, we obtain similar results (NAs:  $\rho = 0.637$ ,  $P = 0.011$ ; Hispanics:  $\rho = 0.096$ ,  $P = 0.254$ ).

We have not controlled for the bias in the estimation of individual genetic admixture away from the extremes. This bias is a consequence of the boundaries on a proportion (0% and 100%), since ancestry levels less than 0% and greater than 100% are undefined. However, as the effect of this bias is to generally make estimates slightly smaller when they are large and slightly larger when they are small, our results are conservative with respect to this bias.

### Skin pigmentation

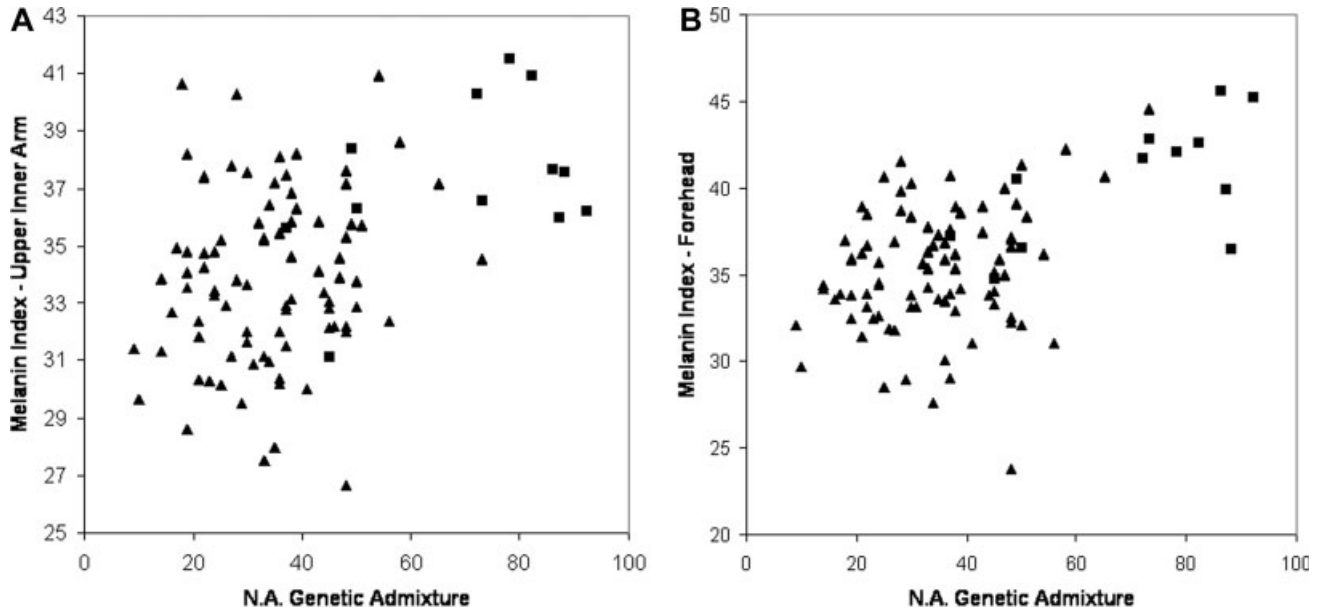
We next examined the relationship between NA genetic admixture and skin pigmentation on the upper inner arm and the forehead. Across the whole sample, a modest proportion of the variation in skin darkness in both areas is explained by variation in NA ancestry (see Fig. 3a,b). When self-identified Hispanics are analyzed separately, the relationship is significant only on the forehead, and not on the upper inner arm. When self-identified NAs are analyzed separately, the relationship is not statistically significant on the forehead ( $\rho = 0.448$ ,  $P = 0.145$ ) or on the upper inner arm ( $\rho = 0.203$ ,  $P = 0.527$ ). Using the dihybrid admixture model, we obtain similar results, except that the relationship for forehead pigmentation is statistically significant ( $\rho = 0.622$ ,  $P = 0.031$ ) among NAs. Insofar as the upper inner arm more directly reveals constitutive (innate) pigmentation, although the forehead also reveals facultative pigmentation (tanning), we suspect that both constitutive and facultative pigmentation increases with NA admixture among Hispanics.

### DISCUSSION

This study suggests that self-identified Hispanics and NAs in New Mexico underestimate their degree of admixture with other groups. This may be due to an evolved tendency in all humans to emphasize a unitary group identity, strongly defined by shared ancestry (Johnson et al., 1987; Salmon, 1998; Jones, 2000; Gil-White, 1999, 2001). In particular, a significant proportion (26%) of self-identified Hispanics claim zero NA ancestry, but virtually all have some NA admixture. The average Hispanic New Mexican in our sample underestimates their level of NA ancestry by about 20%. Conversely, many self-identified NAs claim zero European ancestry, but most have a significant amount of European admixture. Moreover, among self-identified Hispanics, there is no relationship between self-estimated NA ancestry and NA genetic admixture.

There may be several reasons for such discrepancies between genetic admixture and self-identified ethnicity, self-described ethnicity, and self-estimated admixture. First, the genetic admixture analysis has some limitations, such as a large confidence interval, and the inherent problem of using allele frequencies of modern-day parental populations, as well as using Spanish individuals as the parental population representing European admixture. This may present difficulties if a significant portion of the European ancestry among New Mexicans is from regions in Europe other than Spain. There is a strong relationship ( $\rho = 0.93$ ,  $P < 0.001$ , average  $\delta = 0.08$ ) between the allele frequencies of the AIMs between Spanish and the other European regions (Germany and England) for which there is frequency data on these AIMs.

A second reason may be that across all human societies, cultural values associated with "ethnicity" tend to de-emphasize any ancestral ambiguity, possibly related



**Fig. 3.** Skin pigmentation as a function of NA genetic admixture using the trihybrid model (a) Upper-inner arm—across whole sample:  $\rho = 0.29$ ,  $P = 0.003$  (males:  $\rho = 0.449$ ,  $P = 0.011$ ; females:  $\rho = 0.219$ ,  $P = 0.063$ ); Hispanics only:  $\rho = 0.16$ ;  $P = 0.142$  (males:  $\rho = 0.18$ ,  $P = 0.395$ ); females:  $\rho = 0.15$ ,  $P = 0.248$ ); (b) Forehead—across whole sample:  $\rho = 0.36$ ,  $P < 0.001$ , (males:  $\rho = 0.54$ ,  $P = 0.002$ ; females:  $\rho = 0.26$ ,  $P = 0.028$ ); Hispanics only:  $\rho = 0.22$ ,  $P = 0.04$  (males:  $\rho = 0.35$ ,  $P = 0.096$ ; females:  $\rho = 0.14$ ,  $P = 0.264$ ).

to an evolved cognition that “essentializes” ethnic categories (Gil-White, 2001). In New Mexico, this may be due to the lingering effects of political and social pressures on ethnic identity stemming from New Mexico’s bid for statehood in the late nineteenth century (Nieto-Phillips, 2004). In addition, the sample used in this study may not be representative of the New Mexico population. Since it is composed of undergraduate students at a state university, it may overrepresent individuals of higher socioeconomic status (SES). Previous studies have found a negative relationship between income and NA ancestry (Lisker et al., 2004; Martinez-Marignac et al., 2007). Given our sample of college students may have higher than average SES levels, it may be that the individuals in this study may have a lower degree of NA admixture compared with the overall New Mexico population. Self-identified Hispanics of higher SES may also be less likely to acknowledge any mixed ancestry because of sociocultural, political, and historical factors unique to New Mexico, and also because of a tendency that might be universal in humans to associate oneself more with the ethnic group that has the higher SES in a given society. Similarly, self-identified NAs may attach different social significance to their European versus NA ancestry. Our finding that NAs overestimate their degree of NA ancestry may be consistent with how strongly NAs value their ethnic identity, and perhaps with the economic, social, and/or educational benefit that such ancestry confers in contemporary US society.

A third reason may involve the way in which ethnicity/ancestry is socially perceived and determined via phenotypes, language, and culture. In many societies, “skin color” is used as an indicator of social status and/or ethnicity, but skin pigmentation is correlated only very weakly with NA genetic admixture in the present sample. Consequently, individuals may inaccurately assess their degree of genetic admixture if they rely

partly on their skin color or on the cultural use of the label “Spanish”. Furthermore, given the possible “ambiguity and uncertainty” of calling one-self Spanish (Gonzales, 1997), it may be that this label is used by individuals, simply because it makes sense, given their phenotype, or given their family, community, or cultural background (Montalvo and Codina, 2001). However, the relationship between facial appearance and genetic admixture may be tenuous (Parra et al., 2003; Suarez-Kurtz et al., 2007).

With respect to skin color, across the entire sample, there is a statistically significant relationship between pigmentation and NA genetic admixture. Within self-identified Hispanics, the relationship is statistically significant for forehead pigmentation, but not significant for upper-inner arm pigmentation. The stronger relationship for forehead skin color may be due to two factors. There may be an amplification of this relationship for forehead skin color because of a correlation between the tanning response and constitutive pigmentation (Wagner et al., 2002). Alternatively, it may be that those individuals with a higher degree of NA admixture not only have a higher constitutive pigmentation, but are also more exposed to sunlight, possibly because of socioeconomic factors. For example, Martinez-Marignac et al. (2007) and Lisker et al. (2004) find a positive association between SES and European genetic admixture among Mexican residents, perhaps owing to differences in rural versus urban historical residence and educational attainment.

The discrepancies between subjective ethnic identity and genetic admixture have several implications. One implication concerns genetically informed health care. Different populations have differences in disease risk (Risch et al., 2002; Bonilla et al., 2004b; Permutt et al., 2005); however, mismatches have been observed between how individuals identify themselves and how they are

identified by others, and consequently how they may be classified genetically (Barnholtz-Sloan et al., 2005; Borrell, 2005; Gomez et al., 2005; Sinha et al., 2006). Our research suggests that a patient's self-estimated admixture—and even their objectively measured skin pigmentation—may be a rather inaccurate index of a patient's actual genetic ancestry. On the other hand, labels of sub-ethnic identity within self-identified Hispanics are fairly well associated with different levels of NA admixture. Those who described themselves as “half White, half Hispanic” show lower NA admixture (22.6%) than those who described themselves as “Spanish” (26.0%), followed in order by “Hispanic” (30.1%), “Latino/Chicano” (36.7%), “Mexican American” (39.4%), and “Mexican” (46.9%), according to the trihybrid analysis. Reliance on more detailed patient reports of ethnicity or AIM genotyping may therefore provide better information concerning that patient's ancestry proportions, allowing health workers to better diagnose, treat, or assess risk for some conditions (Bamshad, 2005; Barnholtz-Sloan et al., 2005; Sinha et al., 2006; Torres and Kittles, 2007). More knowledge of both the genetic patterns (Bamshad, 2005; Torres and Kittles, 2007) and environmental patterns (Brunner, 2007; Bates et al., 2008) associated with group membership can bring substantial benefits to patients in the coming era of genetically informed medicine within multicultural societies.

A second implication concerns the social reification of race as a biological construct. Although some have expressed concern regarding the potentially negative social consequences arising from genetic ancestry information (Davis, 2000; Foster and Sharp, 2002), this study suggests that individuals can be more “essentialist” in their view of ancestry than is shown by the genetic admixture results. The tendency to essentialize groups and to avoid ambiguity in how one identifies their ancestry, as this study may show, likely has deep evolutionary roots in the behavioral ecology of humans, for whom strong ingroup cooperation and outgroup competition were strong multilevel selective forces (Bowles, 2006). This is apparent today in the context of current-day conflicts (McKie, 2001), as well as ethnographic and sociological accounts (Sherif et al., 1981; Hogg and Abrams, 1987; Gil-White, 1999, 2001). In addition, although skin color is commonly used as a marker of ethnicity in many societies, this study shows it to be of minimal reliability in predicting admixture proportions, especially within ethnic groups. Ancestry as determined by AIMs is a more accurate indicator of ancestry than skin color, because AIMs sample many loci, whereas skin color may depend on just a handful of loci (McEvoy et al., 2006; Stokowski et al., 2007; Sulem et al., 2007).

A third implication concerns how the heterogeneity in genetic admixture proportions within a socially defined group (Hispanics) can be explained, to some extent, by social subdivisions based on geography (Mexican, Spanish), genealogy (Half-Hispanic-Half White), or culture and ethnicity (Spanish). The Spanish ethnic label that has existed in New Mexico for over 100 years has gained much attention from sociologists and anthropologists (Gonzales, 1985, 1997; Nieto-Phillips, 1997; Montgomery, 2002). The results from this study show that those who identify as Spanish have significantly lower levels of NA admixture than Mexicans, thus partly corroborating the sociocultural reasoning for the development of this ethnic term in the nineteenth century (Nieto-Phillips, 1997).

## CONCLUSIONS

This study demonstrates that the relationship between genetic and socially based measures of identification can be confounded by historical, cultural, social, and perhaps phenotypic factors. In the cases where biogeographical ancestry is known to confer a genetically based disease risk, it particularly highlights the potential difficulty in accurately assessing disease risk based on broadly based ethnic identification or on physical appearance. Genetic measures of ancestry or collection of more detailed ethnic histories are likely to be more accurate ways of assessing disease risk and prognosis. In addition, this study demonstrates that more research is warranted on how individuals perceive their admixture proportions and whether increased knowledge about personal genetic ancestry reduces or exacerbates people's ethnocentrism, by either undermining the culturally reified distinctions between human populations, or by giving scientific legitimacy to social and biological prejudices.

## ACKNOWLEDGMENTS

The authors thank all the subjects who participated in this study. The authors also thank Drs. Indrani Halder and Keith Hunley as well as several anonymous reviewers for their help with statistical analyses and revisions of earlier drafts of this manuscript.

## LITERATURE CITED

- Bamshad M. 2005. Genetic influences on health: does race matter? *JAMA* 294:937–946.
- Barbujani G. 2005. Human races: classifying people vs. understanding. *Curr Genomics* 6:215–226.
- Barnholtz-Sloan JS, Chakraborty R, Sellers TA, Schwartz AG. 2005. Examining population stratification via individual ancestry estimates versus self-reported race. *Cancer Epidemiol Biomarkers Prev* 14:1545–1551.
- Bates LM, Acevedo-Garcia D, Alegria M, Krieger N. 2008. Immigration and generational trends in body mass index and obesity in the United States: results of the national Latino and Asian American survey, 2002–2003. *Am J Public Health* 98:70–77.
- Bertoni B, Budowle B, Sans M, Barton SA, Chakraborty R. 2003. Admixture in Hispanics: distribution of ancestral population contributions in the continental United States. *Hum Biol* 75:1–11.
- Bonilla C, Parra EJ, Pfaff CL, Dios S, Marshall JA, Hamman RF, Ferrell RE, Hoggart CL, McKeigue PM, Shriver MD. 2004a. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann Hum Genet* 68(Part 2):139–153.
- Bonilla C, Shriver MD, Parra EJ, Jones A, Fernandez JR. 2004b. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. *Hum Genet* 115:57–68.
- Borrell LN. 2005. Racial identity among Hispanics: implications for health and well-being. *Am J Public Health* 95:379–381.
- Bowles S. 2006. Group competition, reproductive leveling, and the evolution of human altruism. *Science* 314:1569–1572.
- Brunner E. 2007. Biology and health inequality. *PLoS Biol* 5:e267.
- Brutsaert TD, Parra EJ, Shriver MD, Gamboa A, Palacios J-A, Rivera M, Rodriguez I, León-Velarde F. 2004. Effects of birthplace and individual genetic admixture on lung volume and exercise phenotypes of Peruvian Quechua. *Am J Phys Anthropol* 123:390–398.
- Choudhry S, Coyle NE, Tang H, Salari K, Lind D, Clark SL, Tsai HJ, Naqvi M, Phong A, Ung N, Matallana H, Avila PC, Casal J, Torres A, Nazario S, Castro R, Battle NC, Perez-

- Stable EJ, Kwok PY, Sheppard D, Shriver MD, Rodriguez-Cintron W, Risch N, Ziv E, Burchard EG. 2006. Population stratification confounds genetic association studies among Latinos. *Hum Genet* 118:652–664.
- Choudhry S, Seibold MA, Borrell LN, Tang H, Serebrisky D, Chapela R, Rodriguez-Santana JR, Avila PC, Ziv E, Rodriguez-Cintron W, Risch NJ, Burchard EG. 2007. Dissecting complex diseases in complex populations: asthma in Latino Americans. *Proc Am Thorac Soc* 4:226–233.
- Davis DS. 2000. Groups, communities, and contested identities in genetic research. *Hastings Cent Rep* 30:38–45.
- Foster MW, Sharp RR. 2002. Race, ethnicity, and genomics: social classifications as proxies of biological heterogeneity. *Genome Res* 12:844–850.
- Gil-White F. 1999. How thick is blood? The plot thickens: if ethnic actors are primordialists, what remains of the circumstantialist primordialist controversy? *Ethnic Rac Stud* 22:789–820.
- Gil-White F. 2001. Are ethnic groups biological 'species' to the human brain: essentialism in our cognition of some social categories. *Curr Anthropol* 42:515–554.
- Gomez SL, Kelsey JL, Glaser SL, Lee MM, Sidney S. 2005. Inconsistencies between self-reported ethnicity and ethnicity recorded in a health maintenance organization. *Ann Epidemiol* 15(1):71–9.
- Gonzales PB. 1985. The protest function of Spanish-American identity in New Mexico: Southwest Hispanic Research Institute Working Papers Series. Report nr 111, Albuquerque, NM, University of New Mexico.
- Gonzales PB. 1997. The categorical meaning of Spanish American identity among blue-collar New Mexicans, circa 1983. *Hispanic J Behav Sci* 19:123–137.
- Gonzalez-Burchard E, Borrell LN, Choudhry S, Naqvi M, Tsai HJ, Rodriguez-Santana JR, Chapela R, Rogers SD, Mei R, Rodriguez-Cintron W, Arena JF, Kittles R, Perez-Stable EJ, Ziv E, Risch N. 2005. Latino populations: a unique opportunity for the study of race, genetics, and social environment in epidemiological research. *Am J Public Health* 95:2161–2168.
- Gutierrez RA. 1991. When Jesus came, the corn mothers went away. Stanford, California: Stanford University Press.
- Halder I. 2005. Measuring and using individual genomic ancestry to study complex phenotypes. University Park, PA: Pennsylvania State University. 317 p.
- Hanis CL, Chakraborty R, Ferrell RE, Schull WJ. 1986. Individual admixture estimates: disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas. *Am J Phys Anthropol* 70:433–441.
- Hogg MA, Abrams D. 1987. Social identification. New York: Routledge.
- Johnson G, Ratwik S, Sawyer T. 1987. The evocative significance of kin term in patriotic speech. The sociobiology of ethnocentrism. Atlanta, GA: University of Georgia Press.
- Jones D. 2000. Group nepotism and human kinship. *Curr Anthropol* 41:779–809.
- Jorde LB, Wooding SP. 2004. Genetic variation, classification and 'race'. *Nat Genet* 36(11 Suppl):S28–S33.
- Lewontin R. 1972. The apportionment of human diversity. *Evol Biol* 6:391–398.
- Lisker R, Malacara A, Ramirez E, Mutchinick. 2004. Admixture estimates in a Mexican populations stratified by socioeconomic status. *Am J Phys Anthropol* 38(Suppl):136.
- Long JC. 1991. The genetic structure of admixed populations. *Genetics* 127:417–28.
- Luizon MR, Mendes-Junior CT, De Oliveira SF, Simoes AL. 2008. Ancestry informative markers in Amerindians from Brazilian Amazon. *Am J Hum Biol* 20:86–90.
- Marrero AR, Das Neves Leite FP, De Almeida Carvalho B, Peres LM, Kommers TC, Da Cruz IM, Salzano FM, Ruiz-Linares A, Da Silva Junior WA, Bortolini MC. 2005. Heterogeneity of the genome ancestry of individuals classified as White in the state of Rio Grande do Sul, Brazil. *Am J Hum Biol* 17:496–506.
- Martinez-Marignac VL, Valladares A, Cameron E, Chan A, Perera A, Globus-Goldberg R, Wacher N, Kumate J, McKeigue P, O'Donnell D, Shriver MD, Cruz M, Parra EJ. 2007. Admixture in Mexico City: implications for admixture mapping of type 2 diabetes genetic risk factors. *Hum Genet* 120:807–819.
- McEvoy B, Beleza S, Shriver MD. 2006. The genetic architecture of normal variation in human pigmentation: an evolutionary perspective and model. *Hum Mol Genet* 15(Suppl 2):R176–R181.
- McKie R. November 25, 2001. Journal axes gene research on Jews and Palestinians. *The Observer*.
- Montalvo FF, Codina E. 2001. Skin color and Latinos in the United States. *Ethnicities* 1:321–341.
- Montgomery C. 2002. The Spanish redemption. Berkeley, California: University of California Press.
- Nieto-Phillips JM. 1997. "No Other Blood": history, language, and Spanish American ethnic identity in New Mexico, 1880s–1920s. Los Angeles: UMI Company.
- Nieto-Phillips JM. 2004. The language of blood: The making of Spanish American identity in New Mexico, 1880s–1930s. Albuquerque: University of New Mexico Press.
- Parra EJ, Kittles RA, Argyropoulos G, Pfaff CL, Hiester K, Bonilla C, Sylvester N, Parrish-Gause D, Garvey WT, Jin L, McKeigue PM, Kamboh MI, Ferrell RE, Pollitzer WS, Shriver MD. 2001. Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *Am J Phys Anthropol* 114:18–29.
- Parra EJ, Kittles RA, Shriver MD. 2004. Implications of correlations between skin color and genetic ancestry for biomedical research. *Nat Genet* 36(11 Suppl):S54–S60.
- Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD. 1998. Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet* 63:1839–1851.
- Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM, Pena SDJ. 2003. Color and genomic ancestry in Brazilians. *PNAS* 100:177–182.
- Permutt MA, Wasson J, Cox N. 2005. Genetic epidemiology of diabetes. *J Clin Invest* 115:1431–1439.
- Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. 2001. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet* 68:198–207.
- Risch N, Burchard E, Ziv E, Tang H. 2002. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol* 3:comment 2007. p. 1–12.
- Salmon C. 1998. The evocative nature of kin terminology in political rhetoric. *Politics Life Sci* 17:51–57.
- Sherif M, Harvey O, White BJ, Hood WR, Sherif C. 1981. Intergroup conflict and cooperation: The Robber's cave experiment. Norman, OK: University of Oklahoma Press.
- Shriver MD, Parra EJ. 2000. Comparison of narrow-band reflectance spectroscopy and tristimulus colorimetry for measurements of skin and hair color in persons of different biological ancestry. *Am J Phys Anthropol* 112:17–27.
- Shriver MD, Parra EJ, Dios S, Bonilla C, Norton H, Jovel C, Pfaff C, Jones C, Massac A, Cameron N, Baron A, Jackson T, Argyropoulos G, Jin L, Hoggart CJ, McKeigue PM, Kittles R. 2003. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet* 112:387–399.
- Shriver MD, Smith MW, Jin L, Marcini A, Akey JM, Deka R, Ferrell RE. 1997. Ethnic-affiliation estimation by use of population-specific DNA markers. *Am J Hum Genet* 60:957–964.
- Sinha M, Larkin EK, Elston RC, Redline S. 2006. Self-reported race and genetic admixture. *N Engl J Med* 354:421–422.
- Stokowski RP, Pant PV, Dadd T, Fereday A, Hinds DA, Jarman C, Filsell W, Ginger RS, Green MR, van der Ouderaa FJ, Cox DR. 2007. A genomewide association study of skin pigmentation in a South Asian population. *Am J Hum Genet* 81:1119–1132.
- Suarez-Kurtz G, Vargens DD, Struchiner CJ, Bastos-Rodrigues L, Pena SD. 2007. Self-reported skin color, genomic ancestry and the distribution of GST polymorphisms. *Pharmacogenet Genomics* 17:765–771.
- Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Magnusson KP, Manolescu A, Karason A, Palsson A,

- Thorleifsson G, Jakobsdottir M, Steinberg S, Pálsson S, Jonason F, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Kiemenev LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K. 2007. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet* 39:1443–1452.
- Sweeney C, Wolff RK, Byers T, Baumgartner KB, Giuliano AR, Herrick JS, Murtaugh MA, Samowitz WS, Slattery ML. 2007. Genetic admixture among Hispanics and candidate gene polymorphisms: potential for confounding in a breast cancer study? *Cancer Epidemiol Biomarkers Prev* 16:142–150.
- Torres JB, Kittles RA. 2007. The relationship between “race” and genetics in biomedical research. *Curr Hypertens Rep* 9:196–201.
- US Census B. 2000. <http://www.census.gov/prod/2001pubs/c2kbr01-1.pdf>. Accessed on 1 June 2006.
- Wagner JK, Parra EJ, Norton HL, Jovel C, Shriver MD. 2002. Skin responses to ultraviolet radiation: effects of constitutive pigmentation, sex, and ancestry. *Pigment Cell Res* 15:385–390.
- Wang J. 2003. Maximum-likelihood estimation of admixture proportions from genetic data. *Genetics* 164:747–765.
- Williams RC, Knowler WC, Pettitt DJ, Long JC, Rokala DA, Polesky HF, Hackenberg RA, Steinberg AG, Bennett PH. 1992. The magnitude and origin of European-American admixture in the Gila River Indian Community of Arizona: a union of genetics and demography. *Am J Hum Genet* 51:101–110.