Published in final edited form as:

Ann Hum Genet. 2011 July ; 75(4): 529–538. doi:10.1111/j.1469-1809.2011.00655.x.

Wide disparity in genetic admixture among Mexican Americans from San Antonio, TX

Joke Beuten^{1,2,*}, Indrani Halder⁵, Sharon P. Fowler³, Harald H.H. Göring⁶, Ravindranath Duggirala⁶, Rector Arya³, Ian M. Thompson⁴, Robin J. Leach^{1,2,4}, and Donna M. Lehman³

¹ Department of Pediatrics, The University of Texas Health Science Center, San Antonio, Texas

² Department of Cellular and Structural Biology, The University of Texas Health Science Center, San Antonio, Texas

³ Department of Medicine/Clinical Epidemiology, The University of Texas Health Science Center, San Antonio, Texas

⁴ Department of Urology, The University of Texas Health Science Center, San Antonio, Texas

⁵ Cardiovascular Institute, University of Pittsburgh, Pittsburgh, Pennsylvania

⁶ Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Texas

SUMMARY

We studied 706 participants of the San Antonio Family Diabetes Study (SAFDS) and 586 male samples from the San Antonio Center for Biomarkers of Risk of Prostate Cancer (SABOR) and used 64 ancestry informative markers to compare admixture proportions in the two groups. Existence of population substructure was demonstrated by the excess association of unlinked markers. Further, the ancestral proportions differed significantly between the two study groups. In the SAFDS sample, the proportions were estimated at $50.2 \pm 0.6\%$ European, $46.4 \pm 0.6\%$ Native American, and $3.1 \pm 0.2\%$ West African. For the SABOR study sample, the proportions were $58.9 \pm 0.7\%$, $38.2 \pm 0.7\%$ and $2.9 \pm 0.2\%$, respectively. Additionally, in the SAFDS subjects a highly significant negative correlation was found between individual Native American ancestry and skin reflectance ($R^2=0.07$, $p=0.00006$). The correlation was stronger in males than in females but clearly show that ancestry only accounts for a small percentage of the variation in skin color and, conversely, that skin reflectance is not a robust surrogate for genetic admixture. Furthermore, a substantial difference in substructure is present in the two cohorts of Mexican American subjects from the San Antonio area in Texas, which emphasizes that genetic admixture estimates should be accounted for in association studies, even for geographically related subjects.

Keywords

admixture; skin reflectance; Mexican Americans

INTRODUCTION

A method for mapping genes by using admixed populations, first described by Chakraborty and Weiss (Chakraborty and Weiss, 1988), is based on the finding that recently admixed groups have increased levels of linkage disequilibrium (LD) as a result of the intermixing of

*To whom correspondence should be addressed at: Department of Pediatrics, University of Texas Health Science Center at San Antonio/Greehey Children's Cancer Research Institute, 8403 Floyd Curl Drive, San Antonio, TX 78229, Fax : (210)562-9135, beuten@uthscsa.edu.

two or more previously separated populations. In admixed populations this process may result in significant linkage disequilibrium between unlinked marker loci which have large allele frequency differences in the ancestral populations (Pfaff et al., 2001). One such group with a genetically admixed background is Mexican Americans, a subgroup of Hispanics.

Hispanics are the largest and fastest growing minority population in the US and represent approximately 15% of the United States population (<http://www.census.gov>). Although Hispanics are often classified as a single group, the term “Hispanic” describes shared language and cultural heritage, rather than homogeneous bio-geographical ancestry. Self-identified Hispanics have been shown to be genetically heterogeneous with varying proportions of European, Native American and West African admixture (Chakraborty et al., 1999, Hanis et al., 1991, Halder et al., 2008, Bonilla et al., 2004a, Bonilla et al., 2004b). Therefore, studies of complex disorders in Hispanic subjects may be confounded by genetic admixture, in particular if the founding populations differ in trait prevalence (Kittles et al., 2002). Moreover, even among Hispanics who share bio-geographical ancestry, such as Mexican Americans of South Texas, the proportions of genetic admixture may vary widely.

Ancestry informative markers (AIMs) are genetic loci with large frequency differences between populations and used to define the ancestral proportions of admixed groups. Several studies have identified informative sets of AIMs for estimating admixture proportions in Mexican Americans and African Americans (Shriver et al., 2003, Mao et al., 2007, Price et al., 2007, Martinez-Marignac et al., 2007, Halder et al., 2008). Given the genetic heterogeneity observed among Mexican American populations (Tian et al., 2007, Tang et al., 2007, Martinez-Fierro et al., 2009, Martinez-Marignac et al., 2007), we have determined the contribution of three ancestral populations in Mexican Americans from two study cohorts from the San Antonio area using a panel of 64 AIMs. Martinez-Fierro et al. (2009) reported that the minimum number of AIMS markers that accurately distinguished ancestral proportions in their study was 24. Nineteen of these 24 AIMS are included in our set of 64. Additionally, we analyzed the relationship between individual ancestry and skin reflectance that is known to vary between parental groups, and examined possible associations between this trait and marker genotypes. Skin reflectance is a very suitable phenotype for admixture mapping since it shows large variability between parental populations, is measured easily in the inner arm and as such fairly robust to environmental variability.

In this study, we analyzed Mexican Americans from two study cohorts from the San Antonio area, 706 participants of the San Antonio Family Diabetes Study (SAFDS) and 586 participants from the San Antonio Center for Biomarkers of Risk of Prostate Cancer (SABOR) cohort using a panel of 64 AIMs. We determined the contribution of three ancestral populations and the relationship between individual ancestry and skin reflectance in a family based cohort.

MATERIALS AND METHODS

Study Subjects

Subjects from two San Antonio based studies (SAFDS and SABOR) were used in this study.

San Antonio Family Diabetes Study (SAFDS)—This study consists of extended pedigrees and has been described elsewhere (Hunt et al., 2005). Probands for SAFDS were low-income Mexican Americans with type 2 diabetes mellitus from the San Antonio (Texas) geographical area, and all first-, second-, and third-degree relatives of the probands, aged ≥ 18 years, were considered eligible for the study. The 706 SAFDS subjects were distributed across 32 families who had attended either the baseline exam or the second exam; 40.5% were male. Skin reflectance was measured in participants who attended the second exam

using a Photovolt Model 670 portable reflectance spectrophotometer following standard methods (Weiner, 1969). All measurements were taken at the upper inner arm site in order to minimize environmental variation. The tri-stimulus filter amber, sampling the visible wavelength at 600 nm, was used as the skin reflectance measure in this study, which is roughly inversely proportional to melanin content using this measure. Skin reflectance data was available for 180 males (mean 37.9 ± 7.7) and 280 females (mean 40.2 ± 6.8). Ethnicity was self-reported and all subjects identified themselves as Mexican American.

San Antonio Center for Biomarkers of Risk of Prostate Cancer (SABOR)—A second cohort, with male participants from the SABOR cohort, was used for comparison purposes. SABOR is an ongoing study which has been prospectively enrolling healthy male volunteers in San Antonio since 2001. Participants of SABOR receive annual digital rectal examinations and prostate-specific antigen testing. All participants have established US residency and are believed to be long term residents. Ethnicity was self reported. The current cohort is approximately 50% non-Hispanic Caucasian, 36% Hispanic Caucasians and 14% African American. From this cohort, 586 Mexican American individuals were selected who were recruited in a clinic within the San Antonio (Texas) geographical area. No skin reflectance data were available for the SABOR participants. The institutional review board of the University of Texas Health Science Center at San Antonio approved all procedures for SAFDS and SABOR studies, and all subjects gave informed consent.

Description of AIMS and genotyping

In order to estimate admixture proportions in our study cohorts, we used a panel of 64 autosomal AIMS which have large frequency difference in alleles between three parental populations: Europeans, Native Americans and West Africans, as reported previously (Shriver et al., 2003, Martinez-Marignac et al., 2007) and showed an average frequency difference between the parental populations (delta) of 46% for European/Native American populations, 40% for European/West African populations, and 50% for Native American/West African populations. Table 1 lists the 64 markers used in this study, their chromosomal position and allelic frequencies for the specified allele in our two study samples. Frequency data for the ancestral European, Native American and African -populations were obtained from earlier studies (Shriver et al., 2003, Martinez-Marignac et al., 2007) and the public HapMap database (<http://www.hapmap.org/>).

DNA was isolated from participants' whole blood cells or lymphoblastoid cell lines using a QIAamp DNA Blood Maxi Kit (Qiagen, Valencia, CA). Genotyping of the 64 AIMS was performed with the Golden Gate assay of the VeraCode technology using the BeadXpress Reader System according to the manufacturer's protocol (Illumina, San Diego, CA).

Data analysis

The 3LOCUS program was used for LD calculations (Long et al., 1995). The program compares the proportion of markers that are in high LD, which is expected when admixture is present. The LD between physically unlinked markers is computed, and a likelihood ratio statistic (G) for each marker pair is calculated. In an unadmixed sample one can expect about 5% of unlinked marker pairs to show significant LD by chance (at a 0.05 significance level). In a recently admixed sample the proportion of markers with significant LD is inflated. To account for possible LD due to family structure, we used a subset of 95 unrelated founders within the SAFDS sample for 3LOCUS analysis.

Allele frequencies of the AIMS in SAFDS were estimated by a maximum likelihood method using SOLAR, which takes the intra-familial relationships into account (Almasy and Blangero, 1998). Individual admixture was estimated by a maximum likelihood method

implemented in the program Maximum Likelihood Individual Admixture Estimation (MLIAE; <http://izh100.googlepages.com/iae>), as described previously (Halder and Shriver, 2003, Halder et al., 2008). The program requires the subjects' multilocus genotype data and the ancestral allele frequencies from each ancestral population. Briefly, for each genotype, an expression of the likelihood of origin from each ancestral population is derived, based on the allele frequencies in the ancestral populations. The sum of the log-likelihoods for all genotypes for an individual is maximized over the range of possible values of ancestry. For the SAFDs, all samples, i.e. both related and unrelated individuals, were included in the calculation for ancestral proportions. Since each individual is treated as an independent observation when computing admixture estimates, inclusion of other related individuals has no effect on the admixture estimates obtained for each person.

To test for possible associations of the skin reflectance with individual admixture in the SAFDS cohort we used the regression models implemented in the computer program SOLAR, accounting for non-independence of related individuals within families using a kinship variance component that models the additive effects of genes in aggregate (Almasy and Blangero, 1998). The skin reflectance measure was inverse normalized for association analyses. Since the observations are not independent (SAFDS is a family based cohort), the Kuhlback-Leibler R^2 is given, which is the proportion of variance explained by covariates, such as admixture, included in the model. Age and sex were used as covariates in the combined analysis of men and women and only age was used in the analysis stratified by sex.

To evaluate the extent to which the presence of genetic admixture produces false positive and false negative results, we performed a locus-specific association analysis in the SAFDS in which we tested each AIM for association with skin reflectance, using an additive measured genotype approach as implemented in SOLAR (Almasy and Blangero, 1998) with the allele counts serving as the measured genotypes (Boerwinkle et al., 1986). As before, a kinship variance component was included to model the non-independence of relatives, sex and age were included as covariates, and parameter estimates were obtained by maximum likelihood methods. The effect of the genetic admixture on inflation of significance levels in association testing was determined by including the individual ancestral proportions as a covariate in the analysis.

Hardy Weinberg Equilibrium (HWE) was assessed in the SAFDS cohort using SOLAR and in the SABOR cohort using Haploview (<http://www.broad.mit.edu/mpg/haploview/>).

RESULTS

Marker evaluation

The discrepancy rate on 5% duplicate genotyping was $< 0.2\%$ and the call rate was $\geq 97\%$. No Mendelian inconsistencies were observed among the family-based SAFDS samples when using the program SimWalk2 (Sobel et al., 2002). One marker, rs1079598, was not polymorphic in SAFDS and excluded from further analyses. The remaining 63 markers were in HWE when correcting for multiple tests in this group. Markers rs594689 and rs708156 failed amplification in SABOR. Genotype frequencies of 61 markers in SABOR did not deviate from HWE expectations, whereas rs1935946 showed a significant deviation ($p < 0.001$) due to the smaller number of heterozygotes than expected. This marker was excluded from further analysis in SABOR.

Genetic structure and admixture proportions

Since recently admixed populations are known to be genetically stratified (as individual admixture proportions vary in the sample), excess homozygosity or excess heterozygosity is

an indication of presence of genetic stratification. The results from analysis using the 3LOCUS computer program indicate that 10.5% of the unlinked markers pairs in the subset of unrelated individuals of SAFDS and 27.4% of the unlinked marker pairs in SABOR show significant associations (at a 0.05 level), where only 5% are expected to be associated by chance indicating that both populations exhibit significant stratification. Significant LD was found for pairwise combinations of markers located at a distance of >5cM on the same chromosome as well as for two markers located on different chromosomes in both study groups (data not shown).

Based on our set of successfully genotyped AIMs (64 and 62 autosomal AIMs in SAFDS and SABOR, respectively) and a trihybrid model of admixture between Europeans, Native Americans and West Africans, we estimated the composition of the total SAFDS sample as $50.2 \pm 0.6\%$ European, $46.4 \pm 0.6\%$ Native American, and $3.1 \pm 0.2\%$ West African using a maximum likelihood algorithm of the software program MLIAE (Halder et al., 2008, Halder and Shriver, 2003). Figure 1A (left) shows the histograms of the number of individuals falling into each 10% interval of an ancestral population from 0% to 100%. Individual admixture estimates are depicted on the triangle plot are shown in figure 2A (left). When individual admixture was examined, the European and Native American admixture showed a very large variance; individual European admixture ranged from 7.0 to 95% and for the Native American contribution the range was 3.0 to 93% in SAFDS. The distribution did not differ by sex (Table 2). The percent Native American contribution is roughly equivalent to 1 minus the percent European admixture for the majority of subjects.

For the SABOR study sample, the proportions were $58.9 \pm 0.7\%$, $38.2 \pm 0.7\%$ and $2.9 \pm 0.2\%$ for European, Native American, West African ancestry, respectively. Figures 1B and 2B (right) show the histogram of the ancestral proportions per number of individuals and a triangle plot, respectively, for the SABOR sample. Similar to the SAFDS cohort, the European and Native American admixture showed a very large variance with individual European contribution ranging 8.0% to 100% and the Native American contribution ranging from 0.0 to 90%. However, the means of the European and Native American ancestry are substantially different between the SABOR and SAFDS populations (Table 2). As seen in figures 1 and 2, the Native American contribution is lower (and likewise the European contribution is higher) in the SABOR than in the SAFDS.

Individual ancestry and skin pigmentation

Within the SAFDS sample of 460 subjects, skin reflectance exhibits significant residual heritability of $59\% \pm 0.09$ ($p=2.9 \times 10^{-17}$) indicating a genetic basis for differences in skin pigmentation in this cohort. We also observed sex differences in this trait with the females having a higher reflectance, and therefore lighter pigmentation, than males (average reflectance values 37.9 ± 7.1 , males and 40.2 ± 6.8 , females; $p=0.0002$). Although age at the time of skin measurement was not associated with skin reflectance ($p=0.77$), we accounted for the effects of sex and age in all models assessing skin reflectance. Table 3 summarizes the results of association tests for this trait with admixture proportions. Lower skin reflectance (darker skin pigmentation) was significantly associated with higher Native American admixture ($p=0.00006$, $R^2=0.07$), while higher skin reflectance (lighter skin pigmentation) was associated with higher European admixture ($p=0.00004$, $R^2=0.08$), where R^2 is the proportion of the variance due to admixture proportion. West African admixture showed no association with skin color, likely due to the low level of African ancestry in our sample ($p=0.71$, $R^2=0.002$). Figure 3 shows the relationship between individual admixture estimates of Native American proportions and skin reflectance. As can be seen in figure 3, three individuals have unusual high skin reflectance as compared to the remainder of the sample, yet their Native American ancestry estimate is moderate to high. Omission of these outliers had minimal effect on the overall correlation ($R^2=0.073$, $p=0.00003$). We further

noted that there was a significant interaction between sex and ancestry on skin pigmentation ($p=0.047$). The regression on Native American ancestry was more profound in males than in females from the SAFDS ($p=0.0005$ in males, and $p=0.003$ in females) with the proportion of variance in skin reflectance due to Native American admixture proportion higher in males than in females (0.12 and 0.05, respectively). No sex differences were observed with skin reflectance in correlation of West African ancestry. We also determined what proportion of ancestry estimates were accounted for by skin reflectance measures and observed a weak correlation ($R^2=0.01$, $p=0.01$). Skin reflectance data were not available for the SABOR sample.

Locus-specific associations

We tested for locus-specific associations of each AIM with skin reflectance in SAFDS. Since admixture stratification was detected, we performed each test in duplicate. In the first test, only traditional covariates (i.e. age and sex) were included in each model, and in the second test, Native American admixture proportions were included as an added covariate to control for the effects of population stratification. Because of a significant difference in the means of the distribution of skin reflectance between males and females ($p<0.0002$), we included sex as a covariate in the analysis to measure the effect of the genetic markers on skin pigmentation. Our results are summarized in Table 4. Only one AIM (rs1014263) was significantly associated with skin reflectance after correction for 63 independent tests ($p=0.0002$, or roughly 0.013 after Bonferroni testing correction). This association was attenuated after accounting for Native American ancestry ($p=0.0085$) and was no longer significant after correction for multiple testing. In addition, six other AIMs were nominally associated with skin-reflectance, and only three remained nominally significant after adjustment for Native American ancestry. The significance of rs2613964 was stronger after adjustment for admixture. None of these AIMs that are associated map within or close to reported candidate genes for skin pigmentation. The Q-Q plots (Figure 4) confirm the confounding of Native American ancestry; p-values for analyses in which we adjusted only for age and sex but not admixture, show an upward inflation (left). Inclusion of Native American ancestry as an additional covariate in the analysis on the other hand indicates a good fit (right).

DISCUSSION

Studies of common diseases for which risk varies with admixture proportions might lead to false-positive or false-negative association, in particular at loci where allele frequencies vary between subpopulations (Kittles et al., 2002). Therefore knowing the proportion of an individual's ancestry that originated from different populations will help reduce the risk of confounding due to population substructure and admixture and thereby help in identifying true genetic and environmental factors that underlie complex diseases. Our findings of the proportion of Native American ancestry in Mexican Americans from the San Antonio area (40–45%) confirm previous published observations that showed levels between 18–46% in San Antonio, Texas (Relethford et al., 1983), and 40% in northern California (Collins-Schramm et al., 2002). However, lower levels of Native American proportions were reported in Mexicans from Arizona (29%) (Long et al., 1991), Starr County, Texas (30%) (Hanis et al., 1986, Hanis et al., 1991, Cerda-Flores et al., 1992), Colorado (30–34%) (Parra et al., 2004, Bonilla et al., 2004a), and Mexico City (30%) (Martinez-Marignac et al., 2007) and higher levels (56%) of Native American proportions were found in the Mestizo population from Mexico City (Martinez-Fierro et al., 2009). The observed differences between our estimates of Native American ancestral proportion and some previous reports may reflect differences in the geographical origin of the population sampled, the markers that were analyzed, the statistical method, and/or by the number of individuals analyzed. We

found that in our two cohorts the distribution of the European and Native American ancestral estimates differed substantially with less European and more Native American admixture in SAFDS as compared to the SABOR sample. The observed differences between both cohorts could be explained by the time spent in the US. The SABOR participants have US residency which may indicate that they are descendants of immigrants, and not immigrants themselves, and therefore have increased European admixture. In addition, since participants in the SAFDS cohort had type 2 diabetes, this sample may be enriched in individuals with higher Native American ancestry as this disease is more prevalent in indigenous Americans. However, neither type 2 diabetes nor prostate cancer was associated with admixture proportions. Alternatively, socio-economical and/or educational status could also play a role in both inter- and intra-population genetic structure. Indeed, a significant association between European admixture proportion and higher education status has been reported (Florez et al., 2009, Martinez-Marignac et al., 2007). The samples recruited as part of the SAFDS cohort are mainly from the low income area of San Antonio, whereas the SABOR sample covers a more diverse socio-economic group of participants. However, no socio-economic variables were considered in this study and we therefore cannot confirm that social factors are contributing to our findings. Regardless, this observation of a different distribution of the European and Native American ancestral estimates between both cohorts emphasizes that population substructure and admixture should be accounted for in association studies, even for geographically related subjects.

A significant inverse correlation between individual Native American ancestry (or positive with European ancestry) and skin reflectance was observed in the SAFDS. Halder et al. (2008) (Halder et al., 2008) and Bonilla et al. (2004) (Bonilla et al., 2004b) also found an association between skin pigmentation and Native American ancestry, although such correlation was not observed in other studies (Marshall et al., 1993, Bonilla et al., 2004a). The relationship was moderate for such a complex trait as skin pigmentation and only 7% of the variation in skin pigmentation in the SAFDS can be accounted for by the degree of Native American (or conversely, European) ancestry. The majority of the variation in this trait, however, cannot be attributed to this global measure of ancestry and, likewise, the majority of the variation in ancestry proportions cannot be predicted by skin pigmentation. As an extreme example, a number of subjects in this study had high skin reflectance measures but moderate proportions of Native American ancestry. Interestingly, the correlation between individual Native American (or European) ancestry and skin reflectance was more profound in males as compared to females and this in the absence of significant differences between the sexes in average individual ancestry. Moreover, a significant difference in the means of the distribution of skin reflectance was found between males and females of the SAFDS with the men on average darker than females. Although similar pigmentation differences between males and females have been observed (Jablonski and Chaplin, 2000, Shriver et al., 2003, Madrigal and Kelly, 2007), an explanation for this observation remains to be determined and confirmation in a larger sample is warranted.

The confounding effects of genetic structure due to admixture are demonstrated in the regression analyses with skin reflectance using the proportion of individual ancestry as a conditioning variable. Controlling for individual Native American ancestry notably attenuated the association between individual AIMs and skin reflectance. However, some of the variation in the trait could likely be explained by additional factors correlated with ancestry as well as there may be residual stratification that has not been accounted for by the correction of Native American ancestry levels.

A possible limitation of the study is that we used SOLAR for association analysis that does not take into account the uncertainty of the estimation of individual ancestry and thus more accurate results might be obtained when using programs specifically designed for admixture

mapping, such as ADMIXMAP (Hoggart et al., 2003) and STRUCTURE (Pritchard et al., 2000). However, in its current version, ADMIXMAP does not handle family data since it is based on a Bayesian model that uses information from all individuals included in the data set and possible family structures are not accounted for. Families that differ in their admixture proportions as compared to others will thus adversely influence all other estimates. Since the SAFDS cohort consists of pedigrees, we applied a maximum likelihood method from the software program MLIAE to these data to account for non-independence of subjects. Furthermore, this program only requires the marker allele frequency information from the ancestral populations as input as opposed to STRUCTURE for which genotype information is needed. For replication purposes and to be consistent in both cohorts, we also applied MLIAE for the independent males of the SABOR cohort.

Conclusions

In summary, the results from the SAFDS and SABOR samples indicate that a substantial different substructure is present in the two cohorts of Mexican American subjects from the San Antonio area. Our results therefore emphasize that in association studies the risk of confounding due to population substructure and admixture should be accounted for by using individual ancestry estimates as conditioning variable, even for geographically related subjects. We detected a clear association of ancestry levels with skin reflectance, but also observed that skin pigmentation is not a robust surrogate for ancestry estimates and its use as a surrogate can introduce major errors in the data. Further, genetic ancestry estimates only account for a small proportion of the variation in skin color. This study further provides information for a method applicable in other familial studies.

Acknowledgments

The participation and cooperation of all study subjects is gratefully acknowledged. The study could not have been accomplished without the skilled assistance of the SAFDS and SABOR clinical staff. We especially thank Michael P. Stern, Professor Emeritus UTHSCSA, for his guidance and editorial comments. This work was funded in part by NIH grants DK47482, DK70746 and by the San Antonio Cancer Institute, by NCI grant #5U01CA086402 from the Early Detection Research Network of the National Cancer Institute and by American Cancer Society grant #TURSG-03-152-01-CCE.

References

- Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet.* 1998; 62:1198–211. [PubMed: 9545414]
- Boerwinkle E, Chakraborty R, Sing CF. The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann Hum Genet.* 1986; 50:181–94. [PubMed: 3435047]
- Bonilla C, Parra EJ, Pfaff CL, Dios S, Marshall JA, Hamman RF, Ferrell RE, Hoggart CL, McKeigue PM, Shriver MD. Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. *Ann Hum Genet.* 2004a; 68:139–53. [PubMed: 15008793]
- Bonilla C, Shriver MD, Parra EJ, Jones A, Fernandez JR. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York city. *Hum Genet.* 2004b; 115:57–68. [PubMed: 15118905]
- Cerda-Flores RM, Kshatriya GK, Bertin TK, Hewett-Emmett D, Hanis CL, Chakraborty R. Gene diversity and estimation of genetic admixture among Mexican-Americans of Starr County, Texas. *Ann Hum Biol.* 1992; 19:347–60. [PubMed: 1616290]
- Chakraborty BM, Fernandez-Esquer ME, Chakraborty R. Is being Hispanic a risk factor for non-insulin dependent diabetes mellitus (NIDDM)? *Ethn Dis.* 1999; 9:278–83. [PubMed: 10421091]
- Chakraborty R, Weiss KM. Admixture as a tool for finding linked genes and detecting that difference from allelic association between loci. *Proc Natl Acad Sci U S A.* 1988; 85:9119–23. [PubMed: 3194414]

- Collins-Schramm HE, Phillips CM, Operario DJ, Lee JS, Weber JL, Hanson RL, Knowler WC, Cooper R, Li H, Seldin MF. Ethnic-difference markers for use in mapping by admixture linkage disequilibrium. *Am J Hum Genet.* 2002; 70:737–50. [PubMed: 11845411]
- Florez JC, Price AL, Campbell D, Riba L, Parra MV, Yu F, Duque C, Saxena R, Gallego N, Tello-Ruiz M, Franco L, Rodriguez-Torres M, Villegas A, Bedoya G, Aguilar-Salinas CA, Tusie-Luna MT, Ruiz-Linares A, Reich D. Strong association of socioeconomic status with genetic ancestry in Latinos: implications for admixture studies of type 2 diabetes. *Diabetologia.* 2009; 52:1528–36. [PubMed: 19526211]
- Halder I, Shriver M, Thomas M, Fernandez JR, Frudakis T. A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications. *Hum Mutat.* 2008; 29:648–58. [PubMed: 18286470]
- Halder I, Shriver MD. Measuring and using admixture to study the genetics of complex diseases. *Hum Genomics.* 2003; 1:52–62. [PubMed: 15601533]
- Hanis CL, Chakraborty R, Ferrell RE, Schull WJ. Individual admixture estimates: disease associations and individual risk of diabetes and gallbladder disease among Mexican-Americans in Starr County, Texas. *Am J Phys Anthropol.* 1986; 70:433–41. [PubMed: 3766713]
- Hanis CL, Hewett-Emmett D, Bertin TK, Schull WJ. Origins of U.S. Hispanics. Implications for diabetes. *Diabetes Care.* 1991; 14:618–27. [PubMed: 1914811]
- Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, McKeigue PM. Control of confounding of genetic associations in stratified populations. *Am J Hum Genet.* 2003; 72:1492–1504. [PubMed: 12817591]
- Hunt KJ, Lehman DM, Arya R, Fowler S, Leach RJ, Goring HH, Almasy L, Blangero J, Dyer TD, Duggirala R, Stern MP. Genome-wide linkage analyses of type 2 diabetes in Mexican Americans: the San Antonio Family Diabetes/Gallbladder Study. *Diabetes.* 2005; 54:2655–62. [PubMed: 16123354]
- Jablonski NG, Chaplin G. The evolution of human skin coloration. *J Hum Evol.* 2000; 39:57–106. [PubMed: 10896812]
- Kittles RA, Chen W, Panguluri RK, Ahaghotu C, Jackson A, Adebamowo CA, Griffin R, Williams T, Ukoli F, Adams-Campbell L, Kwagyan J, Isaacs W, Freeman V, Dunston GM. CYP3A4-V and prostate cancer in African Americans: causal or confounding association because of population stratification? *Hum Genet.* 2002; 110:553–60. [PubMed: 12107441]
- Long JC, Williams RC, McAuley JE, Medis R, Partel R, Tregellas WM, South SF, Rea AE, McCormick SB, Iwaniec U. Genetic variation in Arizona Mexican Americans: estimation and interpretation of admixture proportions. *Am J Phys Anthropol.* 1991; 84:141–57. [PubMed: 2021190]
- Long JC, Williams RC, Urbanek M. An E-M algorithm and testing strategy for multiple-locus haplotypes. *Am J Hum Genet.* 1995; 56:799–810. [PubMed: 7887436]
- Madrigal L, Kelly W. Human skin-color sexual dimorphism: a test of the sexual selection hypothesis. *Am J Phys Anthropol.* 2007; 132:470–82. [PubMed: 16685728]
- Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, Leon-Velarde F, Moore LG, Vargas E, McKeigue PM, Shriver MD, Parra EJ. A genomewide admixture mapping panel for Hispanic/Latino populations. *Am J Hum Genet.* 2007; 80:1171–8. [PubMed: 17503334]
- Marshall JA, Hamman RF, Baxter J, Mayer EJ, Fulton DL, Orleans M, Rewers M, Jones RH. Ethnic differences in risk factors associated with the prevalence of non-insulin-dependent diabetes mellitus. The San Luis Valley Diabetes Study. *Am J Epidemiol.* 1993; 137:706–18. [PubMed: 8484362]
- Martinez-Fierro ML, Beuten J, Leach RJ, Parra EJ, Cruz-Lopez M, Rangel-Villalobos H, Riego-Ruiz LR, Ortiz-Lopez R, Martinez-Rodriguez HG, Rojas-Martinez A. Ancestry informative markers and admixture proportions in northeastern Mexico. *J Hum Genet.* 2009; 54:504–9. [PubMed: 19680268]
- Martinez-Marignac VL, Valladares A, Cameron E, Chan A, Perera A, Globus-Goldberg R, Wachter N, Kumate J, McKeigue P, O'Donnell D, Shriver MD, Cruz M, Parra EJ. Admixture in Mexico City: implications for admixture mapping of type 2 diabetes genetic risk factors. *Hum Genet.* 2007; 120:807–19. [PubMed: 17066296]

- Parra EJ, Hoggart CJ, Bonilla C, Dios S, Norris JM, Marshall JA, Hamman RF, Ferrell RE, McKeigue PM, Shriver MD. Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado. *J Med Genet.* 2004; 41:e116. [PubMed: 15520398]
- Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. *Am J Hum Genet.* 2001; 68:198–207. [PubMed: 11112661]
- Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, McDonald GJ, Tandon A, Schirmer C, Neubauer J, Bedoya G, Duque C, Villegas A, Bortolini MC, Salzano FM, Gallo C, Mazzotti G, Tello-Ruiz M, Riba L, Aguilar-Salinas CA, Canizales-Quinteros S, Menjivar M, Klitz W, Henderson B, Haiman CA, Winkler C, Tusie-Luna T, Ruiz-Linares A, Reich D. A genomewide admixture map for Latino populations. *Am J Hum Genet.* 2007; 80:1024–36. [PubMed: 17503322]
- Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics.* 2000; 155:945–59. [PubMed: 10835412]
- Relethford JH, Stern MP, Gaskill SP, Hazuda HP. Social class, admixture, and skin color variation in Mexican-Americans and Anglo-Americans living in San Antonio, Texas. *Am J Phys Anthropol.* 1983; 61:97–102. [PubMed: 6869517]
- 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 RA. Skin pigmentation, biogeographical ancestry and admixture mapping. *Hum Genet.* 2003; 112:387–99. [PubMed: 12579416]
- Sobel E, Papp JC, Lange K. Detection and integration of genotyping errors in statistical genetics. *Am J Hum Genet.* 2002; 70:496–508. [PubMed: 11791215]
- Tang H, Choudhry S, Mei R, Morgan M, Rodriguez-Cintron W, Burchard EG, Risch NJ. Recent genetic selection in the ancestral admixture of Puerto Ricans. *Am J Hum Genet.* 2007; 81:626–33. [PubMed: 17701908]
- Tian C, Hinds DA, Shigeta R, Adler SG, Lee A, Pahl MV, Silva G, Belmont JW, Hanson RL, Knowler WC, Gregersen PK, Ballinger DG, Seldin MF. A genomewide single-nucleotide-polymorphism panel for Mexican American admixture mapping. *Am J Hum Genet.* 2007; 80:1014–23. [PubMed: 17557415]
- Weiner, JaLJA. IBP Handbook No. 9. Oxford: Blackwell Scientific Publication; 1969. Human Biology: A guide to field method.

Figure 1.

Histograms showing the distribution of admixture proportions in SAFDS (A, left) and SABOR (B, right)

Figure 2. Triangle plots with individual admixture estimates obtained using 64 AIMS for the SAFDS (A, left) and 62 AIMS for the SABOR (B, right) samples. Each side of the triangle represents a proportion of 0%, with the point of the triangle opposite that base representing a proportion of 100% for the respective ancestry

Figure 3.
Relationship between individual percentage ancestral estimates for the Native American contribution and skin reflectance

Figure 4. QQ-plots of the p values from measured genotype analysis (MGA) adjusted for age and sex (left) or adjusted for age, sex and Native American ancestry (right)

Table 1

Sixty-four ancestry informative markers with their NCBI reference numbers, chromosomal location, major allele, and allele frequencies in our sample groups

SNP	Chromosomal location	Major Allele	SAFDS*	SABOR
rs4908736	1p36.2	A	0.616	0.593
rs2225251	1p34.1	G	0.553	0.532
rs2479409	1p32.3	G	0.609	0.553
rs725667	1p31.3	G	0.943	0.938
rs963170	1p11.2	A	0.521	0.553
rs2814778	1q23.1	A	0.962	0.954
rs6003	1q31.3	A	0.933	0.916
rs1008984	1q32.1	G	0.560	0.598
rs2065160	1q32.1	G	0.538	0.424
rs1506069	1q42.13	A	0.714	0.736
rs2752	1q42.2	A	0.615	0.567
rs1435090	2p16.3	G	0.575	0.545
rs3287	2p16.2	A	0.761	0.776
rs7595509	2p13.3	G	0.560	0.588
rs6730157	2q21.2	G	0.809	0.782
rs1344870	3p24.3	A	0.550	0.596
rs1465648	3p11.2	G	0.795	0.778
rs2317212	3q12.1	A	0.544	0.437
rs2613964	3q13.2	A	0.579	0.536
rs951784	4q28.1	A	0.582	0.615
rs1112828	4q31.1	C	0.538	0.591
rs2702414	4q34.3	G	0.598	0.697
rs26247	5q21.3	A	0.537	0.566
rs3340	5q33.2	A	0.543	0.562
rs2001144	6q21	A	0.682	0.690
rs1935946	6q22.31	C	0.664	0.514
rs1881826	6q24.3	G	0.551	0.645
rs2396676	7q31.1	G	0.722	0.686
rs2341823	7q32.3	A	0.724	0.741
rs1320892	7q35	A	0.508	0.509
rs285	8p21.3	G	0.545	0.547
rs983271	8p12	A	0.693	1.000
rs1808089	8q21.11	G	0.700	0.640
rs4130405	8q22.2	A	0.536	0.599
rs1928415	9p24.3	G	0.899	0.873
rs2888998	9p13.3	A	0.609	0.655
rs2149589	9q21.2	A	0.688	0.590
rs2695	9q21.31	A	0.521	0.490

SNP	Chromosomal location	Major Allele	SAFDS*	SABOR
rs1980888	9q22.1	G	0.505	0.595
rs1594335	10q22.2	A	0.722	0.747
rs2207782	10q23.1	A	0.666	0.584
rs563654	10q24.1	G	0.611	0.703
rs1891760	10q25.3	A	0.596	0.574
rs1487214	11p15.1	A	0.863	0.885
rs594689	11q13.1	G	0.704	failed
rs1042602	11q14.4	C	0.766	0.735
rs1800498	11q23.2	G	0.671	0.574
rs1079598	11q23.2	A	1.000	0.953
rs5443	12p13.31	G	0.618	0.629
rs708156	12p11.23	G	0.541	failed
rs717091	13q14.12	G	0.711	0.746
rs2078588	13q21.33	C	0.921	0.892
rs1800404	15q13.1	A	0.540	0.578
rs2862	15q13.3	A	0.519	0.593
rs724729	15q15.1	G	0.882	0.867
rs4646	15q21.2	C	0.538	0.581
rs292932	16q22.2	A	0.751	0.727
rs2816	17p13.1	G	0.733	0.658
rs1014263	17q21.1	A	0.550	0.676
rs1464612	18q12.1	A	0.680	0.726
rs1369290	18q22.2	C	0.920	0.926
rs1877751	20q13.32	G	0.615	0.589
rs718387	21q22.12	G	0.815	0.815
rs878825	22q11.21	A	0.526	0.600

* Allele frequencies were estimated using SOLAR while accounting for the relationship among pedigree members.

Table 2

Proportion of ancestry in the SAFDS and SABOR sample

	% European ancestry	% Native American ancestry	% West African ancestry
SAFDS (N=706)	50.2 ± 0.6	46.4 ± 0.6	3.1 ± 0.2
Males (N=286)	50.0 ± 0.9	47.3 ± 1.0	2.6 ± 0.2
Females (N=420)	50.3 ± 0.8	46.4 ± 0.8	3.3 ± 0.2
p value (male/female)	p=0.98	p=0.43	p=0.07
SABOR Males (N=586)	58.9 ± 0.7	38.2 ± 0.7	2.9 ± 0.2
p value (SAFDS/SABOR)	p<0.0001	p<0.0001	p=0.14
p value (SAFDS Males/SABOR)	p<0.0001	p<0.0001	p=0.98

Mean ± Standard Error.

Table 3

Negative correlation of individual Native American ancestry with skin reflectance in the SAFDS sample

Skin reflectance	Native American admixture	
	p	R ² *
All	0.00006	6.8%
Females (n=280)	0.003	5.1%
Males (n=180)	0.0005	11.8%

* Proportion of variance in skin reflectance accounted for by the admixture proportion.

Table 4

AIMs exhibiting nominally significant association with skin reflectance in SAFDs

AIMs	Location	Skin reflectance	
		p *	p **
rs4908736	1p36.2	0.034	0.299
rs1435090	2p16.3	0.008	0.063
rs2613964	3q13.2	0.042	0.008
rs2702414	4q34.3	0.394	0.041
rs983271	8p12	0.002	0.026
rs5443	12p13.31	0.108	0.032
rs717091	13q14.12	0.007	0.011
rs4646	15q21.2	0.017	0.190
rs1014263	17q21.1	0.0002	0.009
rs1369290	18q22.2	0.115	0.048

p values are from measured genotype analysis.

* adjusted for age and sex.

** adjusted for age, sex and Native American ancestry.