ORIGINAL INVESTIGATION

Admixture in Mexico City: implications for admixture mapping of Type 2 diabetes genetic risk factors

Veronica L. Martinez-Marignac · Adan Valladares · Emily Cameron ·
Andrea Chan · Arjuna Perera · Rachel Globus-Goldberg · Niels Wacher ·
Jesús Kumate · Paul McKeigue · David O'Donnell · Mark D. Shriver ·
Miguel Cruz · Esteban J. Parra

Abstract Admixture mapping is a recently developed method for identifying genetic risk factors involved in complex traits or diseases showing prevalence differences between major continental groups. Type 2 diabetes (T2D) is at least twice as prevalent in Native American populations as in populations of European ancestry, so admixture mapping is well suited to study the genetic basis of this complex disease. We have characterized the admixture proportions in a sample of 286 unrelated T2D patients and 275 controls from Mexico City and we discuss the implications of the results for admixture mapping studies. Admixture proportions were estimated using 69 autosomal ancestry-informative markers (AIMs). Maternal and paternal contributions were estimated from geographically informative mtDNA and Y-specific polymorphisms. The average proportions of Native American, European and, West African admixture were estimated as 65, 30, and 5%, respectively. The contributions of Native American ancestors to maternal and paternal lineages were estimated as 90 and 40%, respectively. In a logistic model with higher educational status as dependent variable, the odds ratio for higher educational status associated with an increase from 0 to 1 in European admixture proportions was 9.4 (95%, credible interval 3.8–22.6). This association of socioeconomic status with individual admixture proportion shows that genetic stratification in this population is paralleled, and possibly maintained, by socioeconomic stratification. The effective number of generations back to unadmixed ancestors was 6.7 (95% CI 5.7–8.0), from which we can estimate that genome-wide admixture mapping will require typing about 1,400 evenly distributed AIMs to localize genes underlying disease risk between populations of European and Native American ancestry. Sample sizes of about 2,000 cases will be required to detect any locus that contributes an ancestry risk ratio of at least 1.5.

V. L. Martinez-Marignac · E. Cameron ·
A. Chan · A. Perera · R. Globus-Goldberg ·
E. J. Parra (✉)
Department of Anthropology,
University of Toronto at Mississauga,
3359 Mississauga Rd. Room 4026, South Bldg,
L5L 1C6 Mississauga, ON, Canada
e-mail: eparra@utm.utoronto.ca

A. Valladares · M. Cruz
Biochemistry Research Unit, Hospital de Especialidades,
Centro Médico "Siglo XXI" , Instituto Mexicano del Seguro
Social, Mexico DF, Mexico

N. Wacher
Clinical Epidemiology Research Unit,
Hospital de Especialidades, Centro Médico "Siglo XXI" ,
Instituto Mexicano del Seguro Social,
Mexico DF, Mexico

J. Kumate
Fundación IMSS, Mexico DF, Mexico

P. McKeigue · D. O'Donnell
Conway Institute, University College Dublin,
Dublin, Ireland

M. D. Shriver
Department of Anthropology, Penn State University,
University Park, USA

Introduction

The majority of the contemporary Mexican population consists of mestizos, an admixed group with a genetic

background derived from the original Native American inhabitants of Mexico, the European settlers (primarily from Spain) that arrived after the conquest of Mexico by Cortes in the sixteenth century, and to a lesser extent, West Africans who were brought to Mexico mainly during the sixteenth to eighteenth centuries as a consequence of the slave trade in the Americas (Aguirre et al. 1981; Stavenhagen and Carrasco 1997). Studying admixed populations is relevant from both historical and anthropological points of view, and in this respect the modern tools of molecular biology provide unique insights on population history and migration patterns. Additionally, there has been an increased interest in admixed populations in the biomedical field because it is possible to make use of recent admixture to map genes underlying ethnic variation in disease risk. This approach, known as admixture mapping (AM), is analogous to linkage analysis of an experimental cross between inbred strains (Halder and Shriver 2003; Hoggart et al. 2004; Patterson et al. 2004; Montana and Pritchard 2004; Nievergelt and Schork 2005; Smith and O'Brien 2005). Populations with a recent history of admixture (eg., less than 20 generations ago), such as many populations in North, Central and South America and the Caribbean are ideally suited for admixture mapping (Hanis et al. 1991; Parra et al. 1998; Bortolini et al. 1999; Mesa et al. 2000). The history of admixture of each population determines the density of the map required as well as the mapping resolution (Hoggart et al. 2004; Patterson et al. 2004). Thus a first step in planning AM studies is to characterize the history and dynamics of admixture in the population under study.

Type 2 diabetes (T2D) is one of the diseases that can be studied using AM. Risk of T2D varies markedly among population groups. In the US, people of European ancestry have a lower T2D risk than Native American, Latinos, and African American populations (Permutt et al. 2005). Native Americans are 2.2 times more likely to have been diagnosed with diabetes as individuals of European ancestry of a similar age (American Diabetes Association). Similarly, individuals of Mexican ancestry are twice as likely as individuals of European ancestry to have T2D. In principle, admixture mapping should be able to localize the genes underlying these prevalence differences (Halder and Shriver 2003). AM offers important advantages over alternative mapping methods. AM (1) does not require recruitment of families with multiple affected members, in contrast with traditional linkage studies; (2) has higher power to detect variants of modest effect than linkage studies; (3) requires far fewer genetic markers (1,500–3,000) than haplotype or direct association studies (300,000–1 million), with a 100-fold reduction in genotyping costs; (4) is not as affected by allelic heterogeneity as other approaches and 5) can be implemented for affected-only designs (Hoggart et al. 2004; Patterson et al. 2004).

The objective of the present study was to define the design requirements for admixture mapping in T2D by measuring the distribution of admixture proportions and dynamics in a sample of T2D patients ($N = 286$) and controls ($N = 275$) from Mexico City. We used a set of 69 highly informative autosomal AIMs to determine Native American, European, and West African genetic contributions in the sample. We also characterized a panel of mtDNA and Y-chromosome markers to evaluate the history of directional mating. Finally, we tested for evidence of stratification in the sample, and estimated the average number of generations since admixture, a parameter with important implications for AM. We show that the Mexican mestizo population is well suited for the identification and further characterization of T2D genetic risk factors using AM.

Materials and methods

Sample of T2D patients and controls

Samples from 561 unrelated individuals from Mexico City were collected between the years 2000 and 2005 by the Biochemistry and Clinical Epidemiology Research Units of the Medical Center "Siglo XXI", which is a major hospital complex that belongs to the Mexican Institute of Social Security. The samples of T2D patients and controls come from the area of Mexico City served by this Medical Center. The controls were healthy blood donors that were invited to participate in a study to identify Type 2 diabetes (T2D) genetic risk factors. These individuals did not have a family history of T2D and their status as controls was confirmed with a glucose tolerance test, according to the ADA (American Diabetes Association) criteria. Recruitment of the controls took place in the Medical Center "Siglo XXI". T2D patients were recruited from the Family Medicine (Primary Care) clinics associated with the Medical Center "Siglo XXI" within the first 2 years of diagnosis, which was made according to the ADA criteria. Patients with known chronic complications, those on insulin treatment or who were taking drugs that grossly altered glucose metabolism (e.g., glucocorticoids) were excluded from the study. The sample consisted of 286 T2D patients (198 females, 88 males) and 275 controls (86 females, 189 males). Data on sex,

age, BMI, and education were also available. Informed consent was obtained from each participant, and the research was approved by the ethical research board ("Comité Local de Investigación") of the Medical Center "Siglo XXI". The Ethics Review Office at the University of Toronto also approved this study. DNA was isolated from whole blood using the QIAamp DNA Blood Maxi Kit.

Autosomal AIMs

We genotyped 69 AIMs in the sample of cases and controls from Mexico City (Table 1). These markers have large frequency differences between populations of Native American, European, and West African ancestry. Twenty one of the 69 markers (MID-575, FY-NULL, F13B, TSC-1102055, WI-11392, WI-16857, WI-11153, SGC-30610, WI-17163, WI-9231, LPL, WI-11909, D11S429, TYR 192, DRD2-Taq D, DRD2-Bcl I, OCA2, WI-14319, CYP19, WI-7423, and MID-93)

have been described previously (Bonilla et al 2004a, b). Parental frequencies for these AIMs are available for several Native American (Nahua and Maya from Mexico, Native Americans from the Southwestern US), European (Germany, Spain), and West African (Nigeria, Sierra Leone, Central African Republic) population samples. Detailed information regarding these markers can be accessed at dbSNP using PSU-ANTH as the submitter handle. About 47 markers (rs2225251, rs725667, rs963170, rs2814778, rs723822, rs1506069, rs1861498, rs1435090, rs1344870, rs768324, rs1465648, rs2317212, rs719776, rs951784, rs1112828, rs1403454, rs1461227, rs2077681, rs1935946, rs1881826, rs2396676, rs2341823, rs1320892, rs983271, rs1373302, rs1808089, rs1987956, rs1928415, rs1980888, rs1327805, rs1594335, rs2207782, rs1891760, rs1487214, rs726391, rs708156, rs717091, rs2078588, rs724729, rs764679, rs292932, rs1074075, rs1369290, rs386569, rs718092, rs718387, and rs878825) were identified as excellent AIMs in a recent study using the Affymetrix GeneChip

Table 1 Allelic frequencies of 69 autosomal AIMs in West African, European, and Native American parental populations

AIM	Chr.	EUR	N-AM	W-AF	AIM	Chr.	EUR	N-AM	W-AF
rs140864 (MID-575)	1	0.007	0.583	0.137	rs1373302	8	0.287	0.921	0.351
rs2225251	1	0.348	0.808	0.955	rs1808089	8	0.417	0.966	0.397
rs725667	1	0.120	0.001	0.708	rs1987956	8	0.968	0.204	0.918
rs963170	1	0.146	0.922	0.001	rs1928415	9	0.817	0.999	0.250
rs2814778 (FY-NULL)	1	0.993	0.999	0.002	rs2695 (WI-11909)	9	0.167	0.964	0.271
rs723822	1	0.083	0.864	0.219	rs1980888	9	0.933	0.052	0.801
rs6003 (F13B)	1	0.104	0.019	0.698	rs1327805	9	0.814	0.087	0.388
rs1008984	1	0.881	0.276	0.340	rs1594335	10	0.700	0.760	0.206
rs2065160 (TSC-1102055)	1	0.079	0.838	0.504	rs2207782	10	0.347	0.948	0.905
rs1506069	1	0.028	0.368	0.927	rs1891760	10	0.379	0.940	0.203
rs2752 (WI-11392)	1	0.563	0.278	0.153	rs1487214	11	0.064	0.161	0.842
rs1861498	2	0.792	0.991	0.116	rs594689 (D11S429)	11	0.440	0.118	0.089
rs1435090	2	0.240	0.847	0.203	rs1042602 (TYR 192)	11	0.485	0.027	0.004
rs3287 (WI-16857)	2	0.786	0.868	0.285	rs1800498 (DRD2-Taq D)	11	0.630	0.077	0.135
rs1344870	3	0.967	0.060	0.941	rs1079598 (DRD2-Bcl I)	11	0.135	0.572	0.063
rs17203 (WI-11153)	3	0.172	0.814	0.794	rs5443	12	0.681	0.741	0.199
rs768324	3	0.043	0.760	0.205	rs726391	12	0.778	0.500	0.056
rs1465648	3	0.783	0.900	0.109	rs708156	12	0.257	0.847	0.809
rs2317212	3	0.922	0.146	0.295	rs717091	13	0.191	0.319	0.779
rs719776	4	0.880	0.852	0.052	rs2078588	13	0.925	0.875	0.023
rs951784	4	0.171	0.702	0.074	rs1800404 (OCA2)	15	0.636	0.491	0.133
rs1112828	4	0.829	0.113	0.940	rs2862 (WI-14319)	15	0.142	0.704	0.382
rs1403454	4	0.139	0.885	0.026	rs724729	15	0.895	0.999	0.115
rs3309 (SGC-30610)	5	0.300	0.711	0.400	rs4646 (CYP19)	15	0.287	0.739	0.321
rs1461227	5	0.111	0.825	0.409	rs764679	16	0.056	0.625	0.187
rs3340 (WI-17163)	5	0.797	0.216	0.920	rs292932	16	0.001	0.727	0.001
rs2077681	6	0.948	0.788	0.178	rs2816 (WI-7423)	17	0.517	0.061	0.001
rs1935946	6	0.309	0.981	0.801	rs1074075	17	0.266	0.008	0.868
rs1881826	6	0.885	0.200	0.794	rs1369290	18	0.071	0.001	0.900
rs2763 (WI-9231)	7	0.185	0.557	0.138	rs386569	19	0.256	0.948	0.147
rs2396676	7	0.779	0.575	0.129	rs718092	20	0.153	0.824	0.760
rs2341823	7	0.833	0.575	0.126	rs718387	21	0.906	0.893	0.174
rs1320892	7	0.740	0.105	0.904	rs878825	22	0.854	0.241	0.500
rs285 (LPL)	8	0.494	0.439	0.965	rs16383 (MID-93)	22	0.780	0.082	0.280
rs983271	8	0.163	0.491	0.531					

Mapping 10K array (Shriver et al. 2005). In this study we included several parental populations, including Nahua from Mexico, Spanish from Valencia and Mende from Sierra Leone. The usefulness of these AIMs was further confirmed for all but three markers (rs1461227, rs2078588, rs292932) by analyzing independent Nahua, Spanish and West African samples. Finally, two markers were selected (rs1008984, rs5443) based on information previously available in the literature. Table 1 provides information about the AIMs analyzed in this study. The average frequency difference between the parental populations (delta) is 44% for European/Native American populations, 41% for European/West African populations, and 51% for Native American/West African populations. The 69 AIMs were genotyped using McSNP (Akey et al. 2001; Ye et al. 2002) or restriction enzyme assays in the molecular anthropology laboratory of the University of Toronto at Mississauga, or alternatively by allelic specific real time PCR and FRET by the company Kbiosciences (Herts, UK) or a modified allele-specific PCR with Universal energy-transfer-labeled primers by the company Prevention Genetics (Marshfield, Wisconsin, USA). We observed an excellent concordance rate for samples analyzed in duplicate using the same or different genotyping strategies (concordance rates between 99.4 and 100%).

Uniparental informative markers

Due to the limited availability of DNA, we carried out the analysis of uniparental informative markers for a subset of the sample. The mtDNA analysis was carried out in 441 samples (males and females), and the Y-chromosome analysis in 201 males. We genotyped four mitochondrial DNA (mtDNA) haplogroups (A, B, C, D) that represent most of the founding Native American mtDNA lineages (Brown et al. 1998). Additionally, individuals who did not belong to the aforementioned haplotypes were genotyped for haplogroup H (~43% in Spain) (Maca-Meyer et al. 2003; Quintana et al. 2004; Achilli et al. 2004; Pereira et al. 2005) and then for the macrohaplogroup L. Macrohaplogroup L harbors an HpaI restriction site at nucleotide 3592 and includes haplogroups L0, L1, and L2, which are present in 70–100% of the individuals in Sub-Saharan African populations (Torroni et al. 1996). Another African-specific haplogroup, L3, defined by the loss of the HpaI 3592 cut site and the presence/absence of other polymorphisms (Bandelt et al. 2001) was not characterized in the present sample. The mitochondrial haplogroups described above were genotyped following protocols described in Bailliet et al. (1994) and Torroni et al. (1996).

To evaluate the paternal Native American contribution to the sample of Mexico City, we genotyped the DYS199 polymorphism, also known as M3 (Underhill et al. 1996; Underhill et al. 2001). The DYS199*T allele characterizes Native American Y-chromosomes and it is present in approximately 62% of the Y chromosomes in Mayas (Vallinoto et al. 1999), 86% in Mixe males, 70% in Mixtecs, and 50% in Zapotecs males (Lell et al. 1997). We used the protocol described previously by Bianchi et al. (1998) to genotype DYS199. The locus was amplified under standard conditions with a modified reverse primer (5′-TAG-GTACCAGCTCTTCCCAATT-3′), containing a T–C base change (bold and underlined) that creates an artificial MfeI restriction site when the DYS199*C allele is present. The products were resolved on a 3% agarose gel and stained with ethidium bromide. We also characterized the marker M170 (an A to T transition), which is present in European populations (the frequency of this lineage in Spain is approximately 10%) but it is absent in other populations (Maca-Meyer et al. 2004; Brion et al. 2004; Flores et al. 2004; Rootsi et al. 2004). The M170 polymorphism was genotyped according to the conditions described by Ye et al. (2002).

Information about the mtDNA and Y-chromosome markers, including primer sequences and PCR conditions, is provided in Table 2.

Statistical analyses

Average admixture proportions, the sum of intensities parameter (equivalent to the average number of generations since the admixture event) and the individual ancestry proportions were estimated using the software ADMIXMAP v3.2 for Windows. This is a general-purpose program for modeling population admixture with genotype and phenotype data, based on a combination of Bayesian and classical methods. If information for a binary trait (such as T2D) is provided, ADMIXMAP fits a logistic regression model of the trait upon individual admixture. Covariates such as sex, age, BMI, and education can be included in this model. Detailed information about this program can be found in Hoggart et al. (2003, 2004). Versions of ADMIXMAP for Windows and Linux, as well as a binary test data sample and a manual are available at http://www.ucd.ie/genepi/software.html. We specified a model with no "dispersion" of allele frequencies, in which the allele frequencies in the unadmixed modern populations sampled (Native American, European, and West African) were assumed to be the same as the corresponding ancestry-specific allele frequencies in

Table 2 Genotyping conditions for the uniparental markers analyzed in the Mexico City sample

Marker	Alleles	Location	5′–3′ forward/ reverse primers	Annealing temperature (°C)	MgCl$_2$ (mM)	Restriction enzyme
DYS199[a]	C/T	Yq11.221	F- TAATCAGTCTCCTCCCAGCA R- TAGGTACCAGCTCTTCCcAATT	58	2.5	Mfe I-
M170	A/C	Yq11.2 DFFRY Exon 8	F-TGTTTTCATATTCTGTGCATT R-GACACAACCCACACTGAAAAAgA	56.4	1.5	Mnl I-
HAP-A	A/G	mtDNA: 663	L- CTCCTCAAAGCAATACACTGA H- CGTGGTGATTTAGAGGGTGAA	63	1.5	Hae III +
HAP-B	CCCCCTCTA/-9bp	mtDNA: 8272	L-ATTCCCCTAAAAATCTTTGA H-TGCTAAGTTAGCTTTACAG	52.4	1.5	n/a 3% NuSieve 1% agarose
HAP-C	G/A	mtDNA: 13259	L- TAGCAGCAGCAGGCAAATCA H- GGTTGGTTGATGCCGATTGT	64	2	Alu I+
HAP-D	A/C	mtDNA: 5176	L- CCAGCACCACGACCCTACTA H- GGGTGATGGTGGCTATGATG	64	2	Alu I-
HAP-H	C/T	mtDNA:7025	L- GTGGCCTGACTGGCATTGTA H- GGATTTTGGCGTAGGTTTGG	65	1.5	Alu I-
HAP-L	C/T	MtDNA:3592	L- AACCCGCCACATCTACCATC H- GGAGGCCTAGGTTGAGGTTG	65	1.5	Hpa I+
Alu I site		MtDNA:10397	L-TAAGTCTGGCCTATGAGTGACT H-GTAAATGCTAGTATAATATTTA	52.5	2	Alu I+

PCR product run directly in a gel of the specified concentration. Lower case letters in the primers sequence indicate engineered sites

n/a not applicable

[a] Chromosomal location as reported by the UC-Santa Cruz database

the admixed population under study. The program uses this information to estimate the allele frequencies from unadmixed and admixed population samples simultaneously, allowing for sampling error. ADMIXMAP implements a diagnostic test for variation of allele frequencies between the unadmixed populations that were sampled to obtain prior parameters and the corresponding ancestry-specific allele frequencies in the admixed sample. This test is based on the posterior predictive check probability, and small posterior predictive check probabilities indicate lack of fit. Additionally, we used ADMIXMAP to test for stratification unexplained by the fitted model, using a test for residual allelic association between unlinked loci (Hoggart et al. 2003). This test was evaluated using reference priors on allele frequencies for models with 1, 2, and 3 subpopulations. Individual ancestry proportions were also independently estimated using a maximum likelihood method (Chakraborty 1986). Departure of genotype frequencies from Hardy–Weinberg proportions was tested using an exact test form the following web site (http://www.ihg.gsf.de/cgi-bin/hw/hwa1.pl). Gametic disequilibrium coefficients for pairs of loci were estimated using the 3LOCUS program, made available to us by Dr. Jeffrey C. Long. The gametic disequilibrium coefficient is the difference between the observed haplotype frequencies, inferred

by means of an expectation maximization method, and the expected frequencies, which are a function of the observed allele frequencies for a pair of markers. Hypothesis testing was performed with the likelihood ratio statistic (G), which has a $\chi 2$ distribution for large sample sizes.

Because the Native American mtDNA haplogroup X is found nearly exclusively in North America (Schurr et al. 2004) the maternal Native American contribution to the Mexico City sample can be directly estimated as the percentage of mtDNA belonging to the founder Native American haplogroups A, B, C, and D. To estimate the maternal European contribution based on the frequency of the H haplogroup, and the paternal Y chromosome contributions to the sample we used Bernstein's equation: $m = pH–p1/p2–p1$, where pH is the allele frequency in the admixed population and $p1$ and $p2$ are the allele frequencies in the parental populations (Bernstein 1931)

Results

Estimation of ancestry using autosomal AIMs

The results for the average admixture proportions and the sum of intensities parameter in the sample

from Mexico City are depicted in Table 3. The average Native American contribution to the sample was estimated to be 65%, the European contribution 30%, and the West African contribution 5%. The test of dispersion indicated a good fit of the data to the allele frequencies of the unadmixed samples that were used to obtain prior parameters (data not shown). The sum of intensities parameter was estimated as 6.7 per morgan, with 95% credible interval 5.7–8.0. Under a model of admixture occurring at a single pulse, the expected value of this parameter is the number of generations that have elapsed since unadmixed ancestors (Hoggart et al. 2004). Figure 1 shows the distribution of posterior means of individual admixture in the sample from Mexico City using a triangular plot. We also estimated individual admixture proportions using a maximum likelihood approach, and there was a high correlation between the ADMIXMAP and maximum likelihood estimates (correlation of Native American ancestry, $R^2 = 0.986$ $P < 0.001$). The average Native American ancestry was higher in the cases than in the controls (65.5, 64.2%, respectively) but the difference is not significant ($P = 0.210$)

We tested departures from Hardy Weinberg proportions using an exact test. Around 8 out of 69 tests (12%) were significant at an alpha level of 0.05. None of the markers show significant departures from HW after applying the Bonferroni correction for multiple testing.
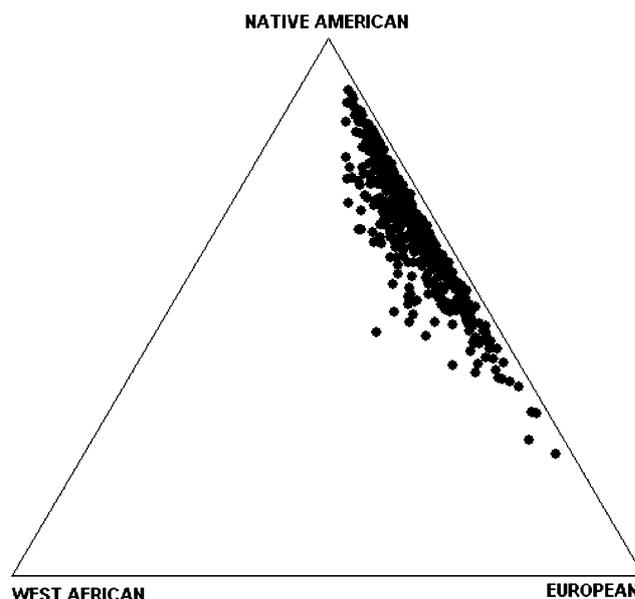

Fig. 1 Triangle plot showing the distribution of individual admixture estimates obtained using 69 autosomal AIMs

Population stratification

The test for residual allelic association between unlinked loci yielded posterior predictive check probabilities of 0.0, 0.52, and 0.50 with models based on 1, 2, and 3 subpopulations, respectively. Typically, values lower than 0.3 indicate strong evidence for residual stratification. Therefore, in the sample from Mexico City a model with two subpopulations was adequate to account for residual allelic associations. An alternative strategy to identify the presence of population stratification is to evaluate the associations between unlinked markers. Out of 1,900 pair wise comparisons between unlinked markers, 442 were significant (23%). Therefore, the percentage of significant associations between unlinked markers is much higher than that expected by chance (5%), indicating that there is a substantial genetic structure in this sample.

Association of Type 2 diabetes with age, sex, BMI, education, and admixture proportion

Table 3 shows the results of the logistic regression analysis with T2D as a dependent variable for a model with age, sex, BMI, education, and individual admixture proportions as predictor variables. The odds ratio for diabetes associated with unit increase in Native American admixture proportion was 1.6, but with a 95% credible interval that overlapped 1 (0.6–4.3). Age, female sex, BMI, and low education were significantly associated with T2D. In a logistic regression analysis

Table 3 Analysis of admixture and estimated odds ratios in logistic regression models with Type 2 diabetes as an outcome variable

Admixture	Median	Mean	Pct2.5	Pct97.5
Sum intensities	6.7	6.7	5.7	8.0
Prop.Nam	0.646	0.646	0.629	0.662
Prop.Eur	0.304	0.304	0.287	0.319
Prop.Afr	0.050	0.050	0.043	0.057
Odds ratios	**Median**	**Mean**	**Pct2.5**	**Pct97.5**
Sex[a]	**0.21**	**0.21**	**0.16**	**0.28**
Age	**1.17**	**1.17**	**1.15**	**1.19**
BMI	**1.05**	**1.05**	**1.02**	**1.09**
Education[b]	**0.57**	**0.57**	**0.50**	**0.66**
Slope Nam	1.62	1.62	0.63	4.35

Sex, age, BMI, and education were included as covariates in the analysis, which was carried out using the program ADMIXMAP. All the covariates were centered on the sample means. Significant covariates are indicated in bold

[a] Female = 1, male = 2

[b] Primary school (primaria) = 1, secundary school (secundaria) = 2, preparatory school (preparatoria) = 3, university degree, and/or postgraduate (universidad y/o postgrado) = 4

with educational status as dependent variable (coded as 1 for university/preparatory level, 0 for no education beyond secondary level), there was a strong association of educational status with European admixture proportion (odds ratio = 9.4, 95% credible interval 3.8–22.6).

Admixture estimates using mitochondrial DNA and Y-specific markers

Table 4 shows the results of the analysis using mtDNA and Y-specific markers. The mtDNA of the majority of the subjects in the Mexico City sample belonged to one of the founder Native American haplogroups: haplogroups A, B, C, and D (41.3, 20.2, 17.2, 7.7, respectively). Furthermore, 13 subjects (2.9%) showed maternal lineages belonging to both A and B haplogroups simultaneously; and one subject (0.2%) showed haplogroups B and D simultaneously. Four individuals (0.9%) showed a triplication of the 9 bp COII/tRNA-Lys intergenic region in the background of a D haplogroup (absence of the *Alu*I 5176 restriction site). Therefore, in this mestizo sample from Mexico City, 90.5% of the individuals have Native American lineages. Only 3.2% of the individuals belonged to the H haplogroup. Using Bernstein's formula with an average frequency of 43% in Spain, we estimated a maternal European contribution of around 7.4% in our sample. Finally, 6.3% of the subjects studied could not be classified into any of the aforementioned haplogroups. We did not observe any African L macrohaplogroup in the sample. Table 4 also depicts the frequencies of two geographically informative Y-specific markers in the sample from Mexico City. About 26.9% of the males harbored the Native American variant DYS199*T.

Table 4 Distribution of mitochondrial DNA haplogroups and Y-specific polymorphisms in the sample from Mexico City

Haplogroup	N	Frequency
A	182	0.413
B	89	0.202
C	76	0.172
D	34	0.077
A/B	13	0.029
B/D	1	0.002
Triplication on D	4	0.009
H	14	0.032
Other	28	0.063
Total	441	1
DYS199*T	54	0.269
M170	12	0.060
Other	135	0.671
Total	201	1

Applying Bernstein's equation using an average DYS199*T frequency of 67% in Native American populations from Mexico, we estimated the Native American paternal contribution to be around 40%. The frequency of the European allele M170 was 6% in this sample. Therefore, using an average frequency of 10% for M170 in Spain, we estimated the paternal European contribution as 60%.

Discussion

Admixture in a sample of T2D patients and controls from Mexico City

Using a panel of highly informative autosomal AIMs and geographically restricted mtDNA and Y-specific polymorphisms, we have characterized the admixture proportions of a sample of T2D patients and controls from Mexico City. The relative Native American and European contributions in the sample were estimated to be 65 and 30%, respectively. There is also a minor West African contribution that was estimated to be 5%. We observed a substantial dispersion in individual admixture estimates, in particular in the Native American and European axes (Fig. 1). It is important to note that the individual admixture estimates are based on a limited number of markers and the variance associated with these estimates is relatively large (>0.1). However, using the Bayesian methods implemented in ADMIXMAP, we can infer the population-level parameters (such as the average admixture proportions and the slope of the regression of education level on individual admixture proportions) accurately in a large sample even where there is not enough information in the marker set for the admixture proportions of each individual to be estimated accurately. It is instructive to compare the average admixture estimates of our sample with previous data from Mexico City, and the rest of Mexico. There have been four previous admixture studies in Mexico City (Tiburcio et al. 1978; Lisker et al. 1986, 1995; Cerda-Flores et al. 2002b). In those studies, the Native American ancestry ranged from 28 to 59%, the European ancestry from 35 to 71%, and the West African ancestry from 1 to 6%. Research on mestizo groups from other states in Mexico also showed a wide dispersion of admixture estimates, with the Native American contribution ranging from 31 to 76%, the European contribution from 16 to 56%, and the West African contribution from 1 to 40% (Crawford et al. 1974, 1976; Lisker et al. 1986, 1988, 1996; Cerda-Flores et al 2002a, b). There are several reasons for the wide

range of ancestry estimates reported in populations throughout Mexico: the type and number of genetic markers used in the different studies, differences in the regional histories of the Mexican states, and differences in the characteristics of the samples (eg., social or economic status). Some of the early admixture studies were based on blood group, serum or red cell enzyme polymorphisms, which are not nearly as informative for inferring admixture proportions as the AIMs used in this study. It is also important to emphasize regional differences in population history within Mexico. For example, the high West African contributions that have been reported in the states located on the east coast of Mexico (eg., Campeche, Yucatan, Tabasco, Veracruz), where West African admixture proportions range between 20 and 40% (Bonilla et al. 2005) are consistent with historical reports indicating a substantial West African presence around the Gulf coast and areas of Southwest Mexico (Oaxaca, Guerrero), regions where the largest Afro-Mexican communities in Mexico are located today (Aguirre et al. 1981).

Additionally, as socioeconomic status is strongly related to individual admixture proportions, we would expect estimates of mean admixture proportions to vary between studies that have sampled different socioeconomic groups. Relethford et al. (1983) estimated admixture using skin reflectance in a sample of Mexican Americans from San Antonio, Texas, and reported that residents of a high-income suburb exhibited the highest European ancestry (82%), while European ancestry levels for residents of a low-income barrio were significantly lower (54%) (Relethford et al. 1983). Our sample of cases and controls corresponds to individuals affiliated with the Mexican Institute of Social Security. This organization serves approximately 50% of the Mexican population, including individuals with different socio-economic levels (Bronfman et al. 1997 and our own data). However, in Mexico the individuals with the highest income often attend private clinics, and the Mexican Institute of Social Security does not cover unemployed persons.

In addition to the evidence from the autosomal AIMs, the information from the uniparentally-transmitted markers (mtDNA and Y-specific markers) offers additional insights on the history of admixture. The mtDNA data indicated that the maternal Native American genetic contribution is approximately 90%, while the Y-chromosome data showed a much lower paternal Native American genetic contribution of around 40% (the average autosomal Native American contribution was 65%). Conversely, the European-specific markers showed the reverse picture, with a European maternal contribution of around 7% and a

paternal contribution of 60% (the average autosomal European contribution was 30%). The widely divergent maternal and paternal estimates clearly indicate that the process of admixture has been sex-biased, with a substantially higher European male contribution than the corresponding female gene flow. This sex-biased contribution has already been described in many other admixture studies throughout the Americas (Merriwether et al. 1997; Dipierri et al. 1998; Batista dos Santos et al. 1999; Sans 2000; Rodriguez-Delfin et al. 2001; Martinez Marignac et al. 2004).

Historical reports indicate that during colonial times Spanish men embarking on the conquest of America commonly practiced unions with Native American women (Herren 1992; Dipierri et al. 1998; Batista dos Santos et al. 1999; Sans 2000; Rodriguez-Delfin et al. 2001; Martinez Marignac et al. 2004; Bonilla et al. 2005). In Mexico, the history of marriages of Native Americans and Spaniards during the three centuries of the colony was recorded primarily by Spanish sources such as marriage records, wills, and lawsuits. These records showed that in most of the cases marriages involved Native American women who married conquerors and *encomenderos* following the royal policies (Grant 1999).

It is of interest to review the distribution of the four Native American haplogroups (A, B, C, D) in our sample, as well as to discuss the presence of individuals harboring mtDNA with more than two haplogroups. In general, the distribution of the four founder haplogroups in the sample from Mexico City is similar to what has been described in other pre-Columbian and contemporary Central American Indigenous people (Torroni et al. 1992; 1994; Green et al. 2000; Gonzalez-Oliver et al. 2001; Bonilla et al. 2005). Haplogroup A is the most common lineage, followed by haplogroups B and C. Haplogroup D is present at very low frequencies. With respect to mtDNA data on mestizo groups, to date there has been only one study reporting data on two populations from north-central Mexico (Green et al. 2000). Similarly to this study, the authors reported a large Native American maternal contribution in the sample (approximately 89%). The distribution of the four founder haplogroups is broadly similar to that of our sample, although in our sample from Mexico City the frequency of haplogroup A is slightly higher (42 vs. 34%) and the frequency of haplogroup B slightly lower (20 vs. 27%) than in the sample from north-central Mexico. Green et al. (2000) reported individuals with both the *Hae*III site at np 663 and the 9bp deletion at the COII/tRNALys intergenic region, which characterize haplogroups A and B, respectively (Green et al. 2000). In our sample we also found 23 individuals harboring this A/B haplogroup. The

occurrence of haplogroups A and B together has been previously reported in Asians (Ballinger et al. 1992; Torroni et al. 1993), Native American populations, including Mayas and Borucas (Bandelt et al. 2001), Mixtecs (Torroni et al. 1994) and Nahuas (Malhi et al. 2003; Bonilla et al. 2005), as well as in Puerto Ricans (Martinez-Cruzado et al. 2001). It is also interesting that in this study, there were subjects whose mtDNAs displayed polymorphisms characteristic of both haplogroups D and B, and we also observed individuals with the triplication of the 9 bp at the COII/tRNALys intergenic region. In contrast with the A/B haplogroup, no D/B haplogroup has been reported previously in Mexican populations. However, this lineage has been reported in other Native American populations (Bandelt et al. 2001; Torroni et al. 1994; Martinez-Cruzado et al. 2001; Malhi et al. 2003), and in Asia (Ballinger et al. 1992; Torroni et al. 1993). In our sample, the 9 bp triplication occurred in four samples under a haplogroup D background. The 9 bp triplication has been described previously in different continental populations and appears to be rare (Shields et al. 1992; Passarino et al. 1993; Merriwether et al. 1995; Alves-Silva et al. 1999). The 9 bp deletion and the 9bp triplication are seen in different mitochondrial backgrounds, often in different continents, and are probably the result of independent deletion/insertion events (Shields et al. 1992; Pasarino et al. 1993; Merriwether et al. 1995; Alves-Silva et al. 1999).

Population stratification

The tests for residual allelic associations implemented in ADMIXMAP indicate that the population is stratified and that a model of admixture between two subpopulations is adequate to account for this stratification. Additionally, 23% of unlinked genetic markers show significant associations, indicating strong evidence of population stratification. Such stratification could be maintained either by continuing gene flow from unadmixed Europeans or Native Americans into this admixed population and/or by social stratification. We were able to generate by simulation a stratified population with similar characteristics to our sample in terms of the admixture proportions (~35%) and the sum of intensities parameter (~7), using a continuous gene flow model with 4% gene flow per generation during 12 generations. Therefore, the stratification that we have observed in the Mexico City sample is compatible with a continuous gene flow model of admixture. However, the strong relationship of educational status to European admixture proportions suggests an alternative mechanism by which stratification could be maintained. In most societies, mating is strongly assortative with respect to socioeconomic status and education (Kalminj 1998) and this pattern has also been described in Mexico (Esteve 2005). Because of the observed association between socioeconomic status and ancestry in Mexico City, assortative mating with respect to socioeconomic status could explain in part the genetic stratification observed in this sample. Irrespective of the mechanism ultimately responsible for the presence of stratification, these results emphasize the importance of controlling for population stratification as a possible confounder in genetic association studies in the Mexican population (Hoggart et al. 2003; Parra et al. 2004). It is important to note that stratification is not a problem for admixture mapping, either using the case-control or affected-only strategies. In the case-control strategy, taking into account the average difference in ancestry between cases, and controls eliminates potential confounding by population stratification due to admixture. In the affected-only test, confounding by stratification is completely controlled because each individual's parental genome is the matched control for that individual.

Implications for applying admixture mapping to Type 2 diabetes

Admixture mapping is a novel gene mapping method that offers great promise to identify genetic risk factors involved in complex traits or diseases showing large differences in prevalence between major continental groups. T2D is at least twice as prevalent in Native American populations as in populations of European ancestry (American Diabetes Association). In our sample from Mexico City, the slope of the logistic regression of T2D upon Native American admixture is positive (odds ratio = 1.6). This is consistent with the available epidemiological data on T2D. However, it is important to mention that the 95% confidence interval overlaps 1 (CI = 0.6–4.3). A larger sample size would be required to explore in more detail the relationship between T2D and ancestry. It is also clear that there is strong confounding by socioeconomic status, which is inversely associated with both diabetes and European admixture proportion.

In order to consider the application of AM in Mexico, it is critical to evaluate the characteristics of the sample. In particular, two important factors to consider are the distribution of individual ancestry and the number of generations since the admixture event. Previous theoretical studies indicated that the power of an AM study is affected by the distribution of admixture proportions, and ideally, the admixture proportions should range

between 10 and 90% (Patterson et al. 2004). In our sample from Mexico City, most of the individuals in the sample are within this range (Native American admixture ranges from 22 to 90%, and the West European contribution between 8 and 75%). Therefore, the distribution of individual ancestry proportions is ideal for an AM study. The second critical parameter to consider when designing an AM study is τ, the sum of intensities parameter. We refer to this parameter as the "effective number of generations back to unadmixed ancestors", as under a model with a single pulse of admixture this is equivalent to the number of generations since the admixture event (Falush et al. 2003). In the sample from Mexico City, τ was estimated as 6.7 per morgan (95% credible interval 5.7–8.0). If the average generation time is 25 years, this is compatible with a single pulse of admixture occurring about 170 years before the present. However, our simulations of a continuous gene flow model show that a sum of intensities of 7 with mean non-Native American admixture proportion 0.35 is also compatible with continuous gene flow at 4% per generation over 12 generations or 300 years. Where mating within the admixed population is assortative with respect to admixture proportions (as is likely where genetic stratification corresponds to social stratification), the sum of intensities parameter will in general underestimate the length of time back to unadmixed ancestors. Our results are thus consistent with the known demographic history of the Mexican population. The sum of intensities parameter is of particular interest for admixture mapping applications because it determines the resolution of AM and the density of markers required to extract information about ancestry at each locus. In admixed samples with a high τ (e.g., resulting from an old admixture event), the mapping resolution is sharper than in admixed samples with low τ, but the density of markers required to extract a given proportion of information is also higher. We estimated τ to be 6.7 in the sample from Mexico City (Table 2). This is equivalent to an effective number of generations since admixture took place between populations of European and Native American ancestry of approximately seven generations. In a previous study of a Hispanic population from San Luis Valley, τ was estimated as 8.1 (Hoggart et al. 2004). For comparison, in two studies of African-American population samples, τ was estimated to be around 6 (Hoggart et al. 2004; Patterson et al. 2004). We have shown that to extract 60% of the information about ancestry across the genome in an initial genome search using a panel of AIMs with an average information content for ancestry (f) of 0.36, an average marker spacing of 3 cM is required for an admixed population where the sum of intensities param-

eter is 6 per morgan (as in African-Americans, (Hoggart et al. 2004). This implies approximately 1,200 AIMs dispersed throughout the genome. In the Mexican mestizo population or the San Luis Valley Hispanic population, where we estimated the sum of intensities parameter to be about 7–8 per morgan, we will require the marker density to be slightly higher: an average spacing of about 2.6–2.3 cM, equivalent to about 1,400–1,600 markers throughout the genome. The benefit of the higher τ in the Hispanic population is that we can also expect to have sharper mapping resolution resulting in a reduced effort in the subsequent fine mapping study. A manuscript published recently by Hoggart et al. (2004) described in detail the calculation of mapping resolution and also the statistical power for AM studies (Hoggart et al. 2004). For the statistical power calculation we assumed admixture proportions of 0.65/0.35, and assumed that regions of putative linkage will be saturated with additional markers to extract nearly all of the information about locus ancestry. Following Hoggart et al. (2004) we estimated the sample sizes required to have 90% power to detect, at $P < 10^{-5}$, loci with different effect sizes. The stringent threshold P value (10^{-5}) allows for the number of independent hypotheses tested in a genome search (or, in Bayesian terms, for the low prior probability of linkage at any given locus). For a rare disease, the required sample sizes to detect a locus with ancestry risk ratios of 1.5, 1.75, 2, and 2.5 are 1,651, 865, 564, and 322, respectively. For a disease like diabetes (prevalence ~10%) the samples required are just slightly larger (around 2,000, ~1,000, ~700, ~400, respectively). Therefore, it will be possible to carry out an AM study in the Mexican population with a relatively small panel of markers (<2,000) and moderate sample sizes (1,000–2,000 individuals). The development of a whole genome panel of AIMs for admixture mapping applications in Hispanics, similar to the panel that was recently described for African Americans (Smith et al. 2004) is a goal within reach. We have recently analyzed four Native American samples and a European sample using the newly developed 500K microarray from Affymetrix (Santa Clara, California) and have identified more than 2,100 AIMs spanning the whole genome (unpublished results). The availability of a whole genome panel will make it possible to apply admixture mapping to identify genetic risk factors playing a role in T2D in the Mexican mestizo population.

Appendix: Simulation of continuous gene flow

To simulate continuous gene flow, we consider a single meiosis in which the paternally derived and maternally derived gametes have admixture proportions μ_1, μ_2, and sum of the intensities of the arrival processes (arrival rates) $\tau 1$, $\tau 2$. The density of ancestry state transitions on a gamete with admixture proportion μ and arrival rate τ is $2\mu(1-\mu)\tau$. The mean density of ancestry state transitions on the two gametes is therefore $\mu_1(1-\mu_1)\tau_1 + \mu_2(1-\mu_2)\tau_2$. The meiosis generates a new gamete with admixture proportion $\mu_{new} = (\mu_1 + \mu_2)\tau 2$. The proportion of meiotic crossovers that generate new transitions of ancestry is $\mu_1(1-\mu_2) + (1-\mu_1)\mu_2$. The arrival rate in the new gamete is therefore

$$\tau_{new} = \frac{\mu_1(1-\mu_1)\tau_1 + \mu_2(1-\mu_2)\tau_2 + \mu_1(1-\mu_2) + (1-\mu_1)\mu_2}{2\mu_{new}(1-\mu_{new})}$$

to simulate gene flow with proportion p of immigrants in each generation into an indigenous population of M gametes, we simulate each generation of new gametes from M meioses based on randomly-sampled pairs of gametes from the indigenous population (which after the first generation includes admixed individuals), and pM meioses based on pairing pM gametes derived from unadmixed immigrants with pM randomly-sampled gametes from the indigenous population. The mean arrival rate is evaluated only over admixed gametes, as unadmixed gametes contribute no information about this parameter. An R script for these simulations is available at http://www.ucd.ie/genepi/software.

References

Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, Cruciani F, Zeviani M, Briem E, Carelli V, Moral P, Dugoujon JM, Roostalu U, Loogvali EL, Kivisild T, Bandelt HJ, Richards M, Villems R, Santachiara-Benerecetti AS, Semino O, Torroni A (2004) The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian Glacial Refuge was a major source for the European gene pool. Am J Hum Genet 75:910–918

Aguirre Beltran G (1981) La población negra de Mexico. In: Secretaría de la Reforma Agrarias (eds) Estudio Etnohistórico Mexico, Centro de Estudios del Agrarismo en Mexico

Akey JM, Sosnoski D, Parra E, Dios S, Heister K, Su B, Bonilla C, Jin L, Shriver MD (2001) Melting curve analysis of SNP's (McSNP): a simple gel-free low-cost approach to SNP genotyping and DNA fragment analysis. Biotechiques 30: 358–367

Alves-Silva J, Guimaraes PE, Rocha J, Pena SD, Prado VF (1999) Identification in Portugal and Brazil of a mtDNA lineage containing a 9-bp triplication of the intergenic COII/tRNALys region. Hum Hered Jan; 49(1):56–58

American Diabetes Association. http://www.diabetes.org/home.jsp

Bailliet G, Rothhammer F, Carnese FR, Bravi CM, Bianchi NO (1994) Founder mitochondrial haplotypes in Amerindian populations. Am J Hum Genet 55: 27–33

Ballinger SW, Schurr TG, Torroni A, Gan YY, Hodge JA, Hassan K, Chen KH, Wallace DC (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient mongoloid migrations. Genetics 130: 139–152

Bandelt HJ, Alves-Silva J, Guimaraes PE, Santos MS, Brehm A, Pereira L, Coppa A, Larruga JM, Rengo C, Scozzari R, Torroni A, Prata MJ, Amorim A, Prado VF, Pena SD (2001) Phylogeography of the human mitochondrial haplogroup L3e: a snapshot of African prehistory and Atlantic slave trade. Ann Hum Genet 65(Pt 6):549–63

Batista dos Santos SE, Rodrigues JD, Ribeiro-dos-Santos AK, Zago MA (1999) Differential contribution of indigenous men and women to the formation of an urban population in the Amazon region as revealed by mtDNA and Y-DNA. Am J Phys Anthropol 109(2):175–180

Bernstein F (1931) Die geographische Verteilung der Blutgruppen und ihre anthropologische Bedeutung. In: comitato Italiano per lo Studio dei Problemi della Populazione, Instituto Poligrafico dello Stato, Roma, pp. 227–243

Bianchi NO, Catanesi CI, Bailliet G, Martinez-Marignac VL, Bravi CM, Vidal-Rioja LB, Herrera RJ, Lopez-Camelo JS (1998) Characterization of ancestral and derived Y-chromosome haplotypes of New World Native populations. Am J Hum Genet 63:1862–1871

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(Pt 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(1):57–68

Bonilla C, Gutierrez G, Parra EJ, Kline C, Shriver MD (2005) Admixture analysis of a rural population of the state of Guerrero, Mexico. Am J Phys Anthropol 128 (4):861–869

Bortolini MC, Da Silva WA Junior W, De Guerra DC, Remonatto G, Mirandola R, Hutz MH, Weimer TA, Silva MC, Zago MA, Salzano FM (1999) African-derived South American populations: a history of symmetrical and asymmetrical matings according to sex revealed by bi- and uniparental genetic markers. Am J Hum Biol 11(4):551–563

Brion M, Quintans B, Zarrabeitia M, Gonzalez-Neira A, Salas A, Lareu V, Tyler-Smith C, Carracedo A (2004) Microgeographical differentiation in Northern Iberia revealed by Y-chromosomal DNA analysis. Gene 329: 17–25

Bronfman M, R Castro E Zúñiga C Miranda J Oviedo (1997) ''Hacemos lo que podemos'': los prestadores de servicios frente al problema de la utilización. Salud Pública de México/vol.39, no.6, noviembre-diciembre

Brown MD, Hosseini SH, Torroni A, Bandelt HJ, Allen JC, Schurr TG, Scozzari R, Cruciani F, Wallace DC (1998)

mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? Am J Hum Genet 63(6):1852–1861

Cerda-Flores RM, Budowle B, Jin L, Barton SA, Deka R, Chakraborty R. 2002a. Maximum likelihood estimates of admixture in Northeastern Mexico using 13 short tandem repeat loci. Am J Hum Biol 14(4):429–439

Cerda-Flores RM, Villalobos-Torres MC, Barrera-Saldana HA, Cortes-Prieto LM, Barajas LO, Rivas F, Carracedo A, Zhong Y, Barton SA, Chakraborty R. 2002b. Genetic admixture in three Mexican Mestizo populations based on D1S80 and HLA-DQA1 loci. Am J Hum Biol 14(2):257–263

Chakraborty R (1986) Gene admixture in human populations: models and predictions. Am J Phys Anthropol 29:1–43

Crawford MH, Leyshon WC, Brown K, Lees F, Taylor L (1974) Human biology in Mexico. II. A comparison of blood group, serum and red cell enzyme frequencies, and genetic distances of the Indian populations of Mexico. Am J Phys Anthropol 41(2):251–268

Crawford MH, Lisker R, Perez-Briceño (1976) Genetic micro-differentiation of two transplanted Tlaxcaltecan populations. In: Crawford M (eds) The Tlaxcaltecans: prehistory, demography, morphology and genetics. Publications in anthropology 7, University of Kansas, Lawrence, p. 169–175

Dipierri JE, Alfaro E, Martinez-Marignac VL, Bailliet G, Bravi CM, Cejas S, Bianchi NO (1998) Paternal directional mating in two Amerindian subpopulations located at different altitudes in North Western Argentina. Hum Biol 70:1001–1010

Esteve A (2005) Tendencias en homogamia educational en México: 1970–2000. Estudios Demográficos y Urbanos 59/20(2)

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure: extensions to linked loci and correlated allele frequencies. Genetics 164:1567–1587

Flores C, Maca-Meyer N, Gonzalez AM, Oefner PJ, Shen P, Perez JA, Rojas A, Larruga JM, Underhill PA (2004) Reduced genetic structure of the Iberian Peninsula revealed by Y-chromosome analysis: implications for population demography. Eur J Hum Genet 12:855–863

González-Oliver A, Márquez-Morfín L, Jiménez JC, Torre-Blanco A (2001) Founding Amerindian mitochondrial DNA lineages in ancient Maya from Xcaret, Quintana Roo. Am J Phys Anthropol 116: 230–235

Grant J (1999) Representing native peoples:recent ethnohistories of colonial Mesoamerica and its frontiers Colonial Latin. Am Rev 8(1):145–152

Green LD, Derr JN, Knight A (2000) MtDNA affinities of the peoples of North-Central Mexico. Am J Hum Genet 66: 989–998

Halder I, Shriver MD (2003) Measuring and using admixture to study the genetics of complex diseases. Hum Genomics 1(1):52–62

Herren R (1992) La conquista erótica de las Indias. Buenos Aires, Argentina. Planeta

Hoggart CJ, Parra EJ, Shriver MD, Bonilla C, Kittles RA, Clayton DG, McKeigue PM (2003) Control of confounding of genetic associations in stratified populations. Am J Hum Genet 72:1492–1504

Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM.2004. Design and analysis of admixture mapping studies. Am J Hum Genet 74(5):965–978

Kalmin M (1998) Intermarriage and homogamy: causes, patterns, trends. Annu Rev Sociol 24:395–421

Lell JT, Brown MD, Schurr TG, Sukernik RI, Starikovskaya YB, Torroni A, Moore LG, Troup GM, Wallace DC (1997) Y chromosome polymorphisms in native American and Siberian populations: identification of Native American Y chromosome haplotypes. Hum Genet 100: 536–543

Lisker R, Babinsky V (1986) Admixture estimates in nine Mexican Indian groups and five east coast localities. Rev Invest Clin 38: 145–149

Lisker R, Pérez Briceño R, Granados J, Babinsky V, Rubens J, Armendares S, Buentello L (1986) Gene frequencies and admixture estimates in a Mexico City population. Am J Phys Anthropol 71: 203–207

Lisker R, Pérez Briceño R, Granados J, Babinsky V (1988) Gene frequencies and admixture estimates in the state of Puebla, Mexico. Am J Phys Anthropol 76(3):331–335

Lisker R, Ramírez E, González-Villalpando C, Stern MP (1995) Racial admixture in a Mestizo population from Mexico City. Am J Hum Biol 7: 213–216

Lisker R, Ramírez E, Babinsky V (1996) Genetic structure of autochthonous populations of Meso-America: Mexico. Hum Biol 68: 395–404

Maca-Meyer N, Sanchez-Velasco P, Flores C, Larruga JM, Gonzalez AM, Oterino A, Leyva-Cobian F (2003) Y chromosome and mitochondrial DNA characterization of Pasiegos, a human isolate from Cantabria (Spain). Ann Hum Genet 67(Pt 4):329–39

Malhi RS, Mortensen HM, Eshleman JA, Kemp BM, Lorenz JG, Kaestle FA, Johnson JR, Gorodezky C, Smith DG (2003) Native American mtDNA prehistory in the American SouthWest. Am J Phys Anthropol 120: 108–124

Martinez-Cruzado JC, Toro-Labrador G, Ho-Fung V, Estevez-Montero MA, Lobaina-Manzanet A, Padovani-Claudio DA, Sanchez-Cruz H, Ortiz-Bermudez P, Sanchez-Crespo A (2001) Mitochondrial DNA analysis reveals substantial Native American ancestry in Puerto Rico. Hum Biol 73: 491–511

Martinez Marignac VL, Bertoni B, Parra EJ, Bianchi NO (2004) Characterization of admixture in an urban sample from Buenos Aires, Argentina, using uniparentally and biparentally inherited genetic markers. Hum Biol 76(4):543–557

McKeigue PM (2005) Prospects for admixture mapping of complex traits. Am J Hum Genet 76(1):1–7

Mesa NR, Mondragon MC, Soto ID, Parra MV, Duque C, Ortiz-Barrientos D, Garcia LF, Velez ID, Bravo ML, Munera JG, Bedoya G, Bortolini MC, Ruiz-Linares A (2000) Autosomal, mtDNA, and Y-chromosome diversity in Amerinds: pre- and post-Columbian patterns of gene flow in South America. Am J Hum Genet 67(5):1277–1286

Merriwether DA, Rothhammer F, Ferrell RE (1995) Distribution of the four founding lineage haplotypes in Native Americans suggests a single wave of migration for the New World. Am J Phys Anthropol 98(4):411–430

Merriwether DA, Huston S, Iyengar S, Hamman R, Norris JM, Shetterly SM, Kamboh MI, Ferrell RE (1997) Mitochondrial versus nuclear admixture estimates demonstrate a past history of directional mating. Am J Phys Anthropol 102: 153–159

Montana G, Pritchard JK (2004) Statistical tests for admixture mapping with case-control and cases-only data. Am J Hum Genet 75(5):771–789

Nievergelt CM, Schork NJ (2005) Admixture mapping as a gene discovery approach for complex human traits and diseases. Curr Hypertens Rep 7(1):31–37

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(6):1839–1851

Parra EJ, Hoggart CJ, Bonilla C, Dios S, Norris JM, Marshall JA, Hamman RF, FerrellP RE, McKeigue M, Shriver MD (2004) 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 41:e116

Passarino G, Semino O, Modiano G, Santachiara-Benerecetti AS (1993) COII/tRNA(Lys) intergenic 9-bp deletion and other mtDNA markers clearly reveal that the Tharus (southern Nepal) have Oriental affinities. Am J Hum Genet 53(3):609–618

Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, Hauser SL, Smith MW, O'Brien SJ, Altshuler D, Daly MJ, Reich D (2004) Methods for high-density admixture mapping of disease genes. Am J Hum Genet 74(5):979–1000

Pereira L, Richards M, Goios A, Alonso A, Albarran C, Garcia O, Behar DM, Golge M, Hatina J, Al-Gazali L, Bradley DG, Macaulay V, Amorim A (2005) High-resolution mtDNA evidence for the late-glacial resettlement of Europe from an Iberian refugium. Genome Res 15(1):19–24

Permutt MA, Wasson J, Cox N (2005) Genetic epidemiology of diabetes. J Clin Invest 115(6):1431–1439

Quintans B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A (2004) Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing. Forensic Sci Int 140:251–257

Relethford JH, Stern MP, Gaskill SP, Hazuda HP (1983) Social class, admixture, and skin color variation in Mexican-Americans and Anglo-Americans living in San Antonio, Texas. Am J Phys Anthropol 61: 97–102

Rodriguez-Delfin LA, Rubin-de-Celis VE, Zago MA (2001) Genetic diversity in an Andean population from Peru and regional migration patterns of Amerindians in South America: data from Y chromosome and mitochondrial DNA. Hum Hered 51: 97–106

Rootsi S, Magri C, Kivisild T, Benuzzi G, Help H, Bermisheva M, Kutuev I, Barac L, Pericic M, Balanovsky O, Pshenichnov A, Dion D, Grobei M, Zhivotovsky LA, Battaglia V, Achilli A, Al-Zahery N, Parik J, King R, Cinnioglu C, Khusnutdinova E, Rudan P, Balanovska E, Scheffrahn W, Simonescu M, Brehm A, Goncalves R, Rosa A, Moisan JP, Chaventre A, Ferak V, Furedi S, Oefner PJ, Shen P, Beckman L, Mikerezi I, Terzic R, Primorac D, Cambon-Thomsen A, Krumina A, Torroni A, Underhill PA, Santachiara-Benerecetti AS, Villems R, Semino O (2004) Phylogeography of Y-chromosome haplogroup I reveals distinct domains of prehistoric gene flow in Europe. Am J Hum Genet 75(1):128–137

Sans M (2000) Admixture studies in Latin America: From the twentieth to the twenty first century. Hum Biol 72(1):155–177

Schurr TG, Sherry ST (2004) Mitochondrial DNA and Y chromosome diversity and the peopling of the Americas: evolutionary and demographic evidence. Am J Hum Biol 16(4):420–439

Shields GF, Hecker K, Voevoda MI, Reed JK (1992) Absence of the Asian-specific region V mitochondrial marker in Native Beringians. Am J Hum Genet 50(4):758–765

Shriver MD, Mei R, Parra EJ, Sonpar V, Halder I, Tishkoff SA, Schurr TG, Zhadanov SI, Osipova LP, Brutsaert TD, Friedlaender J, Jorde LB, Watkins WS, Bamshad MJ, Gutierrez G, Loi H, Matsuzaki H, Kittles RA, Argyropoulos G, Fernandez JR, Akey JM, Jones KW (2005) Large-scale SNP analysis reveals clustered and continuous patterns of human genetic variation. Hum Genomics 2(2):81–89

Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, Waliszewska A, Kessing BD, Malasky MJ, Scafe C, Le E, De Jager PL, Mignault AA, Yi Z, De The G, Essex M, Sankale JL, Moore JH, Poku K, Phair JP, Goedert JJ, Vlahov D, Williams SM, Tishkoff SA, Winkler CA, De La Vega FM, Woodage T, Sninsky JJ, Hafler DA, Altshuler D, Gilbert DA, O'Brien SJ, Reich D (2004) A high-density admixture map for disease gene discovery in African Americans. Am J Hum Genet 74(5):1001–1013

Smith MW, O'Brien SJ (2005) Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. Nat Rev Genet. 6(8):623–632

Stavenhagen R., Carrasco T (1997) La diversidad étnica y cultural. In: Florescano E: El patrimonio nacional de México, 1st edn. Consejo Nacional para la Cultura y las Artes, pp 249–280

Tiburcio V, Romero A, De Garay AL (1978) Gene frequencies and racial admixture in a Mestizo population from Mexico City. Ann Hum Biol 5: 131–136

Torroni A, Schurr TG, Yang CC, Szathmary EJE, Willliams RC, Schanfield MS, Troup GA, Knowler WC, Lawrence DN, Weiss KM, Wallace DC (1992) Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrations. Genetics 130: 153–162

Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, Vullo CM, Wallace DC (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. Am J Hum Genet 53: 563–590

Torroni A, Chen YS, Semino O, Santachiara-Benerecetti AS, Scott CR, Lott MT, Winter M, Wallace DC (1994) MtDNA and Y-chromosome polymorphisms in four Native American populations from southern Mexico. Am J Hum Genet 54: 303–318

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza LL (1996) A pre-Columbian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93: 196–200

Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oefner PJ, Cavalli-Sforza LL (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62

Vallinoto AC, Cayres-Vallinoto IM, Ribeiro Dos Santos AK A, Zago MA, Santos SE, Guerreiro JF (1999) Heterogeneity of Y chromosome markers among Brazilian Amerindians. Am J Hum Biol 11(4):481–487

Ye J, Parra EJ, Sosnoski DM, Hiester K, Underhill PA, Shriver MD (2002) Melting curve SNP (McSNP) genotyping: a useful approach for diallelic genotyping in forensic science. J Forensic Sci 47: 593–600