Geographic Patterns of Genome Admixture in Latin American Mestizos

Sijia Wang[1], Nicolas Ray[2], Winston Rojas[3], Maria V. Parra[3], Gabriel Bedoya[3], Carla Gallo[4], Giovanni Poletti[4], Guido Mazzotti[5‡], Kim Hill[6], Ana M. Hurtado[6], Beatriz Camrena[7], Humberto Nicolini[7], William Klitz[8,9], Ramiro Barrantes[10], Julio A. Molina[11], Nelson B. Freimer[11], Maria Cátira Bortolini[12], Francisco M. Salzano[12], Maria L. Petzl-Erler[13], Luiza T. Tsuneto[13], José E. Dipierri[14], Emma L. Alfaro[14], Graciela Bailliet[15], Nestor O. Bianchi[15], Elena Llop[16], Francisco Rothhammer[16,17], Laurent Excoffier[2], Andrés Ruiz-Linares[1]*

1 The Galton Laboratory, Department of Biology, University College London, London, United Kingdom, 2 Computational and Molecular Population Genetics Laboratory, University of Bern, Bern, Switzerland, 3 Laboratorio de Genética Molecular, Universidad de Antioquia, Medellín, Colombia, 4 Laboratorios de Investigación y Desarrollo, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia, Lima, Perú, 5 Facultad de Medicina, Universidad Peruana Cayetano Heredia, Lima, Perú, 6 Department of Anthropology, University of New Mexico, Albuquerque, New Mexico, United States of America, 7 Departamento de Ciencias Genómicas, Universidad Autónoma de la Ciudad de México, México D.F., México, 8 School of Public Health, University of California Berkeley, Berkeley, California, United States of America, 9 Public Health Institute, Oakland, California, United States of America, 10 Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica, 11 Center for Neurobehavioral Genetics, University of California Los Angeles, Los Angeles, California, United States of America, 12 Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil, 13 Departamento de Genética, Universidade Federal do Paraná, Curitiba, Brazil, 14 Instituto de Biología de la Altura, Facultad de Humanidades y Ciencias Sociales, Universidad Nacional de Jujuy, San Salvador de Jujuy, Argentina, 15 Laboratory of Human Molecular Population Genetics, IMBICE, La Plata, Argentina, 16 Programa de Genética Humana, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile, 17 Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile

Abstract

The large and diverse population of Latin America is potentially a powerful resource for elucidating the genetic basis of complex traits through admixture mapping. However, no genome-wide characterization of admixture across Latin America has yet been attempted. Here, we report an analysis of admixture in thirteen *Mestizo* populations (i.e. in regions of mainly European and Native settlement) from seven countries in Latin America based on data for 678 autosomal and 29 X-chromosome microsatellites. We found extensive variation in Native American and European ancestry (and generally low levels of African ancestry) among populations and individuals, and evidence that admixture across Latin America has often involved predominantly European men and both Native and African women. An admixture analysis allowing for Native American population subdivision revealed a differentiation of the Native American ancestry amongst Mestizos. This observation is consistent with the genetic structure of pre-Columbian populations and with admixture having involved Natives from the area where the Mestizo examined are located. Our findings agree with available information on the demographic history of Latin America and have a number of implications for the design of association studies in population from the region.

Introduction

There is growing interest in the application of admixture mapping to the identification of genes influencing complex traits (including disease) in populations tracing their ancestry to genetically differentiated populations[1–5]. This approach is potentially more powerful and economical than high-density whole genome association studies and should also allow the identification of trait-related genetic variants that are fixed in one of the parental populations. Considerable progress has been made in the application of admixture mapping in African-Americans[6–11]. Similarly, it is hoped that admixture mapping may be a powerful approach for gene identification in populations from Latin America[12], and first generation marker maps for use in these populations have recently been developed[13–15]. Ideally, the application of admixture mapping should build on knowledge regarding the genetic makeup of the admixed population, as well as of the specific ancestral populations that contributed to the admixture. Unfortunately, although it is broadly known that the history of Latin America entailed an extensive admixture of Native

Author Summary

The history of Latin America has entailed a complex process of population mixture between Native and recent immigrants across a vast geographic region. Few details are known about this process or about how it has shaped the genetic makeup of contemporary Latin American populations. To perform a broad exploration of the genetic diversity of Latin America we carried out genome-wide analyses in 13 mestizo populations sampled from 7 countries across the region. We observe a marked variation in ancestry both within and between mestizo populations. This variation in ancestry correlates with pre-Columbian Native population density in the areas examined and with recent patterns of demographic growth of the sites sampled. We also find evidence that the mixture at the origin of these populations involved mainly immigrant European men and Native and African women. Finally, mestizo populations show a differentiated Amerindian genetic background, consistent with a predominantly local Native ancestry. Mestizos thus still reveal the genetic imprint of the pre-Columbian Native American population diversification. Our study helps delineate the genetic landscape of Latin America and has a number of implications for gene identification analyses in populations from the region.

Americans, Europeans and Africans, few details are known about this process or about its genetic correlates[16–19]. Early demographic history data is scant and population genetic studies in the region are so far quite restricted in terms of the number of populations and/or markers that have been examined [20,21]. A genomic survey of admixture in populations across Latin America is therefore of considerable historical interest and is also important for assessing the context in which admixture mapping could be applied in populations from the region.

To help draw a more detailed picture of the genetic landscape of Latin America, here we report genetic diversity and admixture analyses based on microsatellite genome scan data for 249 individuals from 13 urban centers sampled in seven countries across the region (Figure 1 and Table S1). For this study we avoided examining areas of important recent transcontinental immigration (such as the large urban centers of Southern South America) and focused in areas that since colonial times (i.e. prior to the 19th century) have been settled mainly by Natives and Europeans (thus roughly corresponding to the term "Mestizo" populations).

Results/Discussion

We analyzed genotype data for 678 autosomal and 29 X-chromosome microsatellites collected in the Mestizo populations together with similar data available in samples from Europeans, Native Americans and Africans[22,23]. Bar charts summarizing the estimated ancestry proportions of the populations examined are shown in Figure 2 (the corresponding values and standard errors of these estimates are presented in Table S2). The autosomal data indicate substantial variation in Native American ancestry, ranging from ~70% in Salta to ~20% in Rio Grande do Sul (RGS), the Central Valley of Costa Rica (CVCR) and Medellin. African ancestry is low (<5%) in most of the populations examined, although it approaches 10% in Medellin, RGS and Oriente. African ancestry is often accentuated in a few outlier individuals for each population (Figure S1). The observed variation in ancestry is consistent with historical differences in

Native population density and with the extent of past immigration to the regions sampled. The Mestizo with the highest Native ancestry are in areas which historically (and to the present) have had relatively large Native populations: Andean regions (Salta, Huilliche) and meso-America (Mexico City, Oriente), where major pre-Columbian civilizations developed[17,21]. By contrast, the Mestizo with highest European ancestry (CVCR, Medellin and RGS) are from areas with relatively low pre-Columbian Native population density (occupied then by heterogeneous groups of chiefdoms or hunter-gatherers) and where the current Native population is sparse[17,21]. Categorizing the Mestizo examined into three groups, based on the relative pre-Columbian Native population density in the region (Table S1), results in a significant Spearman rank correlation with levels of Native ancestry ($\rho = 0.569$, $P = 0.04$). The highest African ancestry (~10%) occurs in Mestizo in relative proximity to circum-Caribbean areas (Oriente and Medellin) and in Southern Brazil (RGS), and thus at the periphery of regions with large past African immigration. Based on the autosomal data, estimates of the mean time since admixture in the 13 Mestizo populations range between ~6–14 generations (Table S3), in agreement with independent estimates made in some of the populations examined here [14,24] and consistent with the notion that most admixture in these populations is likely to have occurred in colonial times[16]. These age estimates are obtained based on the inferred mean frequency of transitions in ancestry along the genome, under the assumption of a single past admixture event[25]. These estimates should be viewed with caution, as it is doubtful that such as model applies in the populations we examined and the added complication of a non-negligible three-way admixture in some of these populations (Figure 2). The observed variation in the estimated age of admixture in different populations is in fact likely to be influenced by variable levels of historical gene flow in different regions[26]: the relatively more isolated populations (e.g. CVCR)[27,28] tend to show older age estimates, while populations in the vicinity of large local native populations (e.g. Salta, in Northern Argentina)[29] or near areas of recent European immigration (e.g. RGS, in Southern Brazil)[30] show younger estimates, consistent with more recent gene flow and possibly ongoing admixture.

A positive correlation ($\rho = 0.758$, $P < 0.01$) is observed between autosomal heterozygosity and European ancestry (Figure 3). This increase in heterozygosity with higher European ancestry agrees well with expectations, based on the difference in mean diversity of European and Native American populations and their genetic differentiation, as measured by F_{ST} ($\rho = 0.786$, $P < 0.01$; See Methods). No significant correlation is seen between settlement size and genetic diversity or between settlement size and Native American ancestry (results not shown). Large differences in the variation of individual admixture estimates were seen across populations, with the variance in Native American ancestry between individuals ranging from 0.005 in Quetalmahue to 0.07 in Mexico City (Figure 4, Figure S1, and Table S2), an observation consistent with previous studies[31,32].

At the X-chromosome level, the proportions of African and Native American ancestry estimated are usually larger than those based on autosomal markers, with a concomitant reduction in European ancestry (Figure 2, Wilcoxon signed rank test $P = 0.02$). This pattern is consistent with admixture involving predominantly European men and Native women. Such a sex bias in European-Native admixture has been inferred in Mestizo populations mainly based on mtDNA and Y-chromosome polymorphisms[30,33–38] and the data collected here confirm that it is a common phenomenon across Latin America. Interestingly, these data also indicate that a similar sex bias in admixture applies (even more

Figure 1. Approximate Geographic Location of the Mestizo Populations Examined and of the Native American Populations Considered in the Analyses. Mestizo populations are indicated as red triangles with names in bold italic font. The Native populations have been color coded based on their affiliation to one of the main Amerindian linguistic stocks according to the classification of Ruhlen[40]. RGS = Rio Grande do Sul; CVCR = Central Valley of Costa Rica.
doi:10.1371/journal.pgen.1000037.g001

strongly) to the African ancestry in Mestizos: in all the populations examined a higher estimate of African ancestry is observed on the X-chromosome than on autosomes (Figure 2, Wilcoxon signed rank test $P<0.001$). Such a sex bias in African admixture has been inferred African Americans from the US[39] but had not been evidenced in Mestizos. Figure 2 also indicates that the difference in European ancestry between X-chromosome and autosomal markers is positively correlated with the extent of European ancestry of the population ($\rho = 0.736$, $P<0.01$). This suggests that the sex bias of admixture has been more pronounced in areas with lower Native population density, consistent with the observation that Mestizo populations from areas with low Native population density (such as Medellin and CVCR) can have a predominantly

European autosomal background and at the same time an almost exclusively Native American mtDNA ancestry[36]. This pattern could also have been influenced by the collapse of the Native population soon after the establishment of the Mestizo in these regions, and the continuing immigration of European men over several generations[36]. A relatively high sex bias of European/ African admixture in the regions sampled here (possibly associated with a historically low African population density) is consistent with the uniformly higher estimates of African ancestry obtained with the X-chromosome relative to autosomes (Figure 2).

Admixture analyses generally face the difficulty of not knowing with certainty the specific ancestral populations that were involved in the admixture, particularly since such ancestral populations
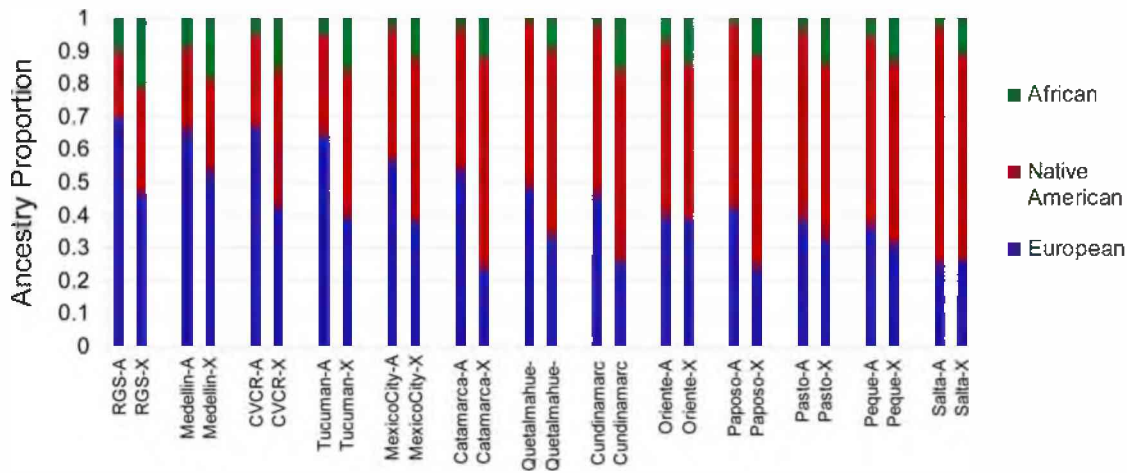
Figure 2. Ancestry Proportions in 13 Mestizo Populations. For each population, proportions estimated with autosomal [-A] and X-chromosome [-X] markers are color-coded on separate bars. The populations have been ordered left to right based on decreasing autosomal European ancestry. The values of these ancestry estimates and their associated standard errors are shown in Table S2. Ancestry was estimated by grouping data for populations sampled in Europe, Africa and Native Americans into three continental population samples. Data for these populations was obtained from the HGDP-CEPH human genome diversity panel database (v 1.0) (http://www.cephb.fr/hgdp-cephdb/) and from Wang et al. (2007)[23].
doi:10.1371/journal.pgen.1000037.g002

might not be available for study or they could have undergone extensive genetic drift. Admixture estimates are therefore usually obtained by pooling data from related putative ancestral population samples, as a way to approximate a "mean" ancestral gene pool. All previous reports of admixture in Latin American populations have therefore pooled population data from African, European and Native American samples into "continental" samples; as done for the analyses discussed above. However, since there is a high level of population structure amongst ancestral Native American populations [23] it is conceivable that the Native component of Mestizos could be genetically differentiated across different geographic regions. We investigated whether it is possible to detect such an underlying genetic differentiation amongst Mestizos through an admixture analysis allowing for a structured ancestral Native American population sample (see Methods). The results from this analysis are not strictly ancestry proportions reflecting an underlying admixture between multiple Native populations. This is particularly so because the proportions

obtained are influenced by the variable degree of genetic relatedness amongst the various Native groups examined. Rather, these proportions reflect the relative genetic similarity of the Native American component in the Mestizo to the Native groups considered in the analyses. Figure 5 shows such a partitioning of the Native American ancestry in Mestizos when admixture is estimated with data from Native Americans subdivided based on linguistic grounds; using the classification of Ruhlen[40] (the corresponding values and errors of these estimates are presented in Table S4). Generally, the native component in the Mestizo shows greatest genetic similarity to Native populations from the linguistic stock which is most widespread in the region where the Mestizo population sampled is located (Figure 1): Central/Northern Amerind in Mexico City and Oriente; Chibchan-Paezan in CVCR, Medellin and Peque; Andean in Pasto, Salta, Catamarca and Quetalmahue. RGS shows no strong similarity to Natives from any linguistic stock but is the Mestizo population with greatest similarity to the Equatorial-Tucano, consistent with RGS
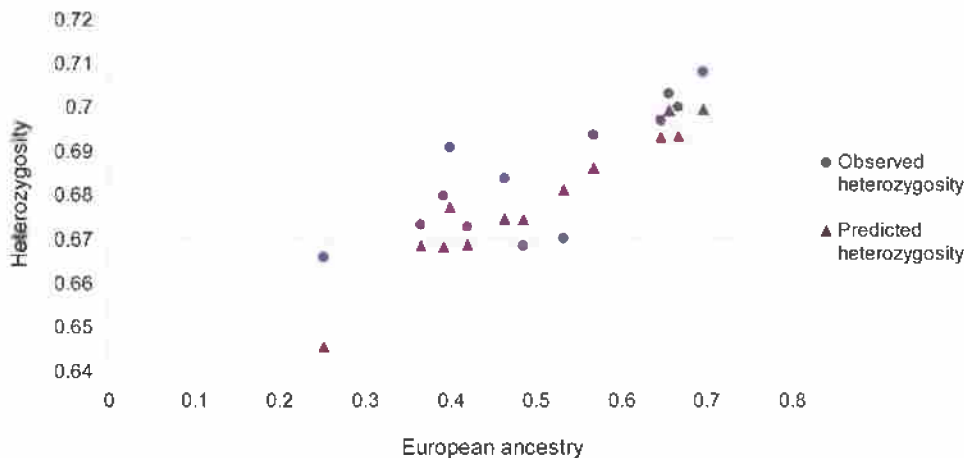
Figure 3. Heterozygosity vs. Proportion of European Ancestry in Mestizo Populations. The heterozygosity predicted from the estimated ancestry of a population was calculated as described in Methods.
doi:10.1371/journal.pgen.1000037.g003
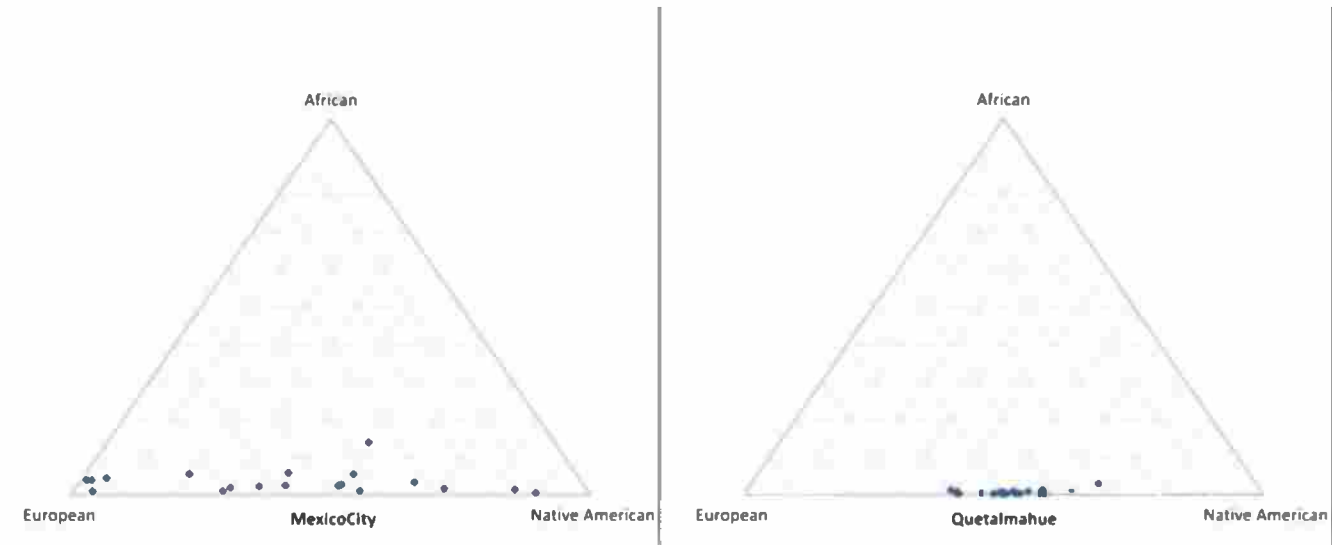
Figure 4. Distribution of Admixture Estimates for Individuals from Mexico City and Quetalmahue. The position of each blue dot on the triangle plot indicates the proportion of European, Native American and African ancestry estimated for each individual in the population. The triangle plots for the other 11 Mestizo populations examined are shown in the Supporting Information (Figure S1).
doi:10.1371/journal.pgen.1000037.g004

being the Easternmost of the Mestizo populations examined (Figure 1). Overall, these observations agree with the expectation that admixture is likely to have involved mainly Natives from the region where the Mestizo populations are located.

Performing a similar admixture analysis but this time considering each individual Native American population as ancestral, revealed several instances of increased similarity between the Native component in the Mestizo and Native populations located in relative geographic proximity (Figure 1, Figure 6 and Table S5). Most notably, Quetalmahue (in Southern Chile) shows a strong

genetic similarity to the Huilliche, a Native population from the vicinity. Also, the population of Paposo in Northern Chile is markedly more similar to the neighboring Aymara than to any other Native American population. The three populations from North West Argentina (Salta, Tucuman and Catamarca) show greatest genetic similarity to the Quechua (sampled in Southern Peru) and the Aymara (sampled in Northern Chile). The population of Pasto in Southern Colombia is most similar to the Inga, a Quechua-speaking population also from Southern Colombia. Peque in North-West Colombia shows greatest
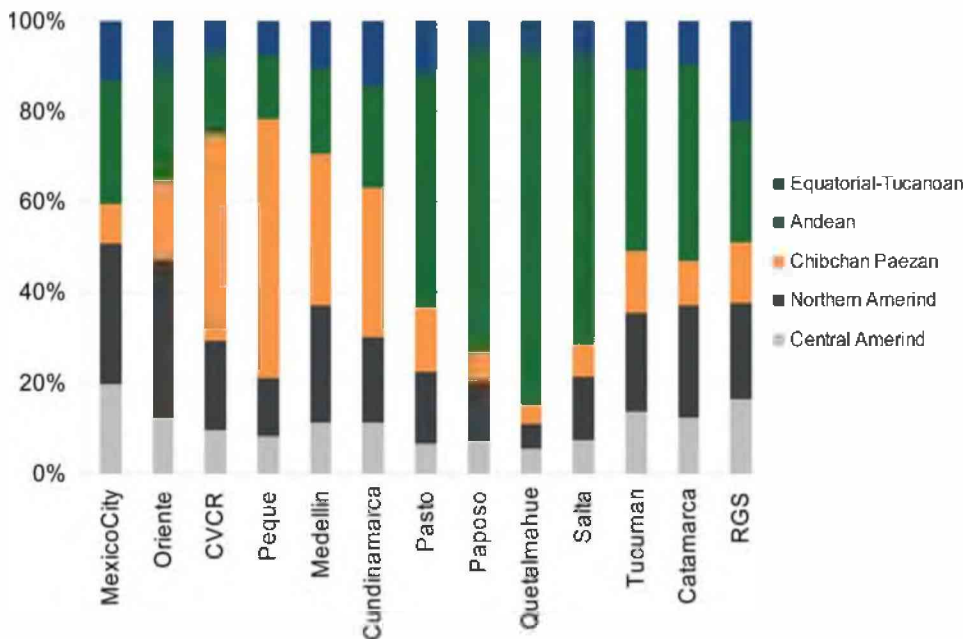
Figure 5. Regional Native American Ancestry of 13 Mestizo Populations Considering the Major Native American Linguistic Stocks. The relative partitioning of the Native American component is shown as the proportion of the colored bar (the European and African components are not shown). Ge-Pano-Carib is not included as it is represented here only by one population, the Kaingang.
doi:10.1371/journal.pgen.1000037.g005
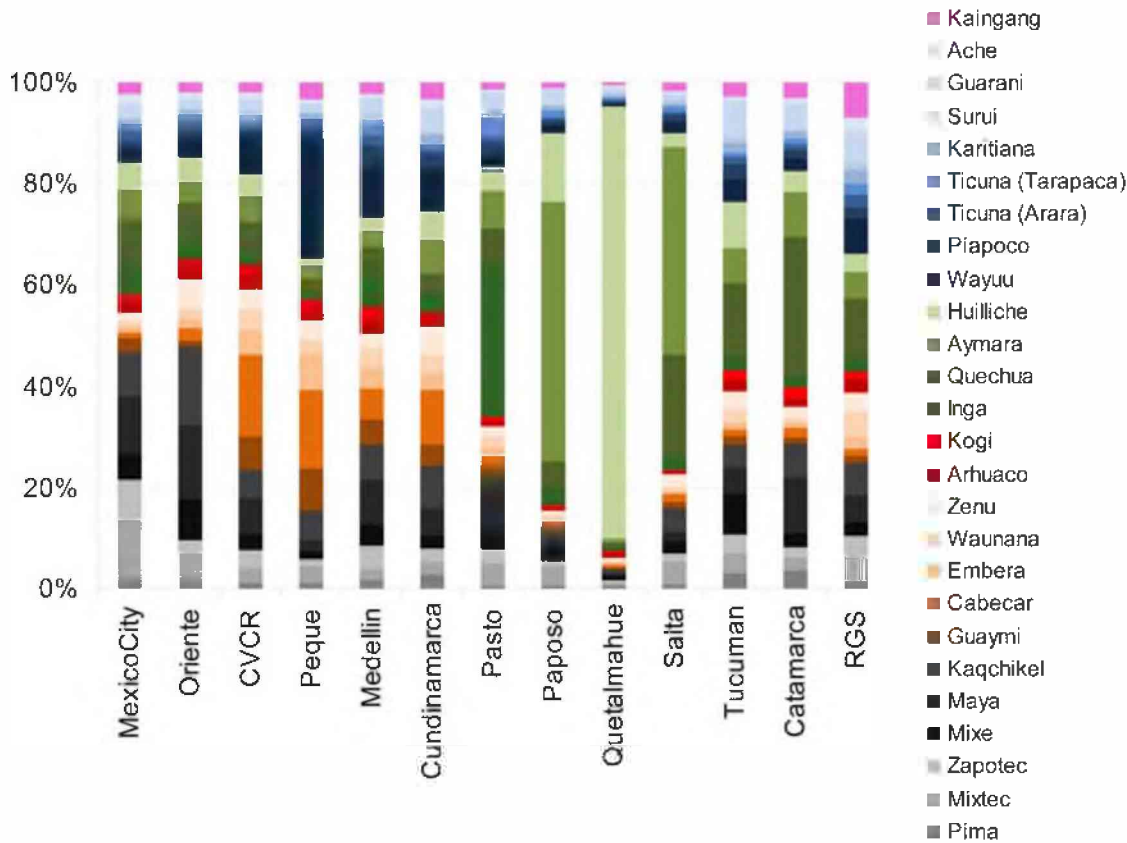
Figure 6. Regional Native American Ancestry of 13 Mestizo Populations Considering each Individual Native Population. The relative partitioning of the Native American component is shown as the proportion of the colored bar (the European and African components are not shown). These are coded in shades of a color corresponding to the main linguistic stocks shown in panel Figure 1 and Figure 5.
doi:10.1371/journal.pgen.1000037.g006

similarity to the Wayuu (sampled in Northern Colombia and genetically closest to Chibchan-speakers, although not classified as Chibchan[23]) and the Cabecar (from Costa Rica, in lower Central America). The Cabecar are also the Native population most similar to the Mestizo population of the Central Valley of Costa Rica (CVCR). Finally, Oriente (in Guatemala) shows greatest genetic similarity to the Maya (sampled in Southern Mexico) and the Kaqchikel (sampled in Guatemala). The populations of Mexico City, Medellin, Cundinamarca and RGS appear to have more heterogeneous Native American ancestries.

The congruence between pre-Columbian genetic structure and the genetic differentiation of the Mestizo is also evidenced in the correlation of the logarithm of the geographic distance between Mestizo and Native populations, and the size of the corresponding ancestry components (as shown in Figure 6). These correlations are negative for all Mestizo populations (Figure 7), in agreement with a stronger genetic affinity of the Mestizo to Native populations that are geographically closer. The correlations are usually higher when considering an effective geographic distance (a distance considering preferential migration along the coastal outline, see Methods), consistent with the influence of the coasts on Paleolithic Native population dispersals[23]. The mean weighted Pearson correlation coefficient over all mestizo populations is -0.481 ($R^2 = 0.232$) for Euclidean distances and -0.570 ($R^2 = 0.325$) for effective distances.

The analysis of admixture considering a structured ancestral Native American population also suggests that a stronger regional ancestry is present in the smaller urban centers sampled. The variance of the estimated Native ancestry components is negatively correlated with the logarithm of population size (Table S1), both in the linguistic-based analysis (Figure 5, $\rho = -0.611$, $P<0.05$) and in the population-based analysis (Figure 6, $\rho = -0.661$, $P<0.05$). The more heterogeneous Native American ancestry of larger urban agglomerations is consistent with them having attracted immigrants from relatively distant areas, thus potentially tracing their ancestry to various, differentiated Native groups. By contrast, smaller urban centers appear to trace their ancestry to fairly defined Native groups, with subsequent maintenance of greater isolation than larger agglomerations. This genetic pattern agrees with demographic data showing that the expansion of major cities in Latin America has often been driven by regional immigration rather than by internal growth[41].

The large variation amongst the Mestizo examined in the mean Native American ancestry and in individual admixture proportions, and the regionally differentiated Native American ancestry, testify to the marked genetic heterogeneity of Latin American populations. These observations have a number of implications for the application of admixture mapping in the region. The large variation in mean Native American ancestry between populations implies that the power of admixture mapping will vary considerably in studies targeting different geographic areas[14]. The differentiated Native Ancestry of Mestizos will affect the informativeness of admixture maps across populations and could result in an increase of false positives when admixture mapping is
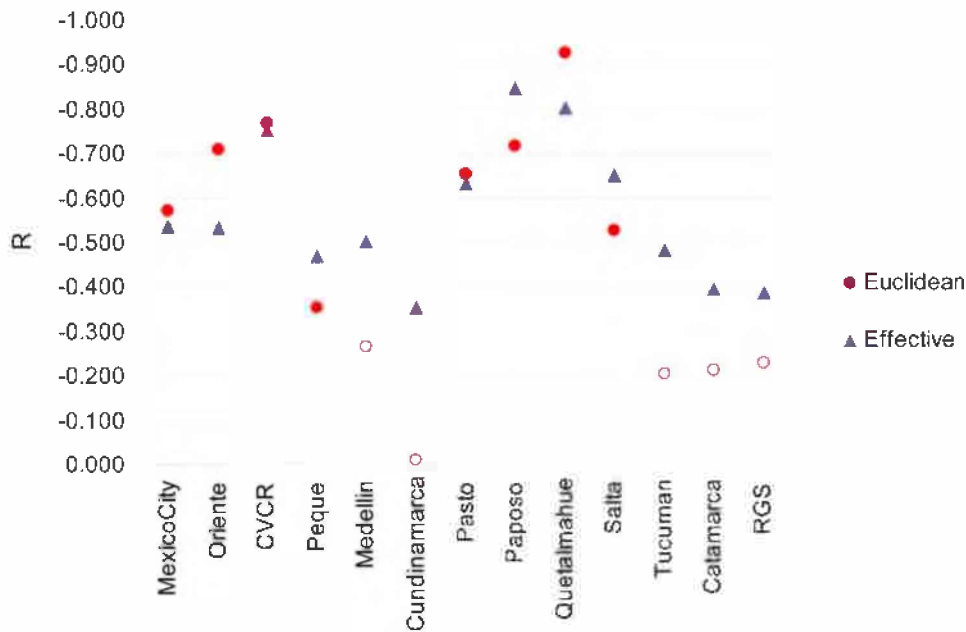
Figure 7. Correlation between Geographic Distance and Estimated Ancestry Components. For each Mestizo population, the Pearson correlation coefficients were calculated between the logarithm of the geographic distance (Euclidean and effective, see methods) separating the Mestizo and Native populations, and the estimated ancestry components (from Figure 6). Correlation coefficients with associated P-values <0.05 are shown as filled symbols.
doi:10.1371/journal.pgen.1000037.g007

attempted populations other than the one from which informative markers were selected. Ideally, admixture maps should therefore be developed for each Mestizo population studied. An alternative would be to select markers for mapping based on their lack of differentiation across Native American populations. Our results also show that mean African ancestry in Mestizo populations is typically low ($<10\%$). This reduces the potential complexity of an extensive three way admixture and confirms that admixture mapping in these populations should be feasible within the two-population admixture framework usually considered[24,42]. Mapping in Mestizos should thus be practical with marker maps that mainly distinguish Native from European ancestry (or Native from non-Native ancestry), possibly supplemented by the exclusion of outlier individuals showing a marked increase in African ancestry. It is likely, however, that areas where historically there has been substantial African immigration (e.g. circum-Caribbean areas) will show higher levels of African admixture and represent additional challenges for admixture mapping. Finally, individual admixture estimates can vary markedly in certain Mestizo populations, particularly in large urban agglomerations such as Mexico City (Figure 4). These populations could be particularly useful for evaluating the effect of ancestry on phenotype, an important initial step prior to admixture mapping of genes influencing such phenotypic variation.

In conclusion, this initial genome-wide analysis of admixture across Latin America has revealed a hitherto undetected differentiation of the Native American ancestry in Mestizos. This fact, together with the extensive variation observed in rates of admixture across populations, and sometimes also between individuals within populations, needs to be considered when designing admixture mapping studies in specific Latin American populations. Despite these complications, we anticipate that admixture mapping in Mestizos should prove a fruitful strategy for analyzing the genetic basis of phenotypic traits, including disease, differing between Native Americans and Europeans.

Methods

Population Samples

A total of 249 unrelated individuals from 13 Mestizo populations were examined (Figure 1 and Table S1). The individuals studied were not selected based on any specific phenotype and no ethnic identification was attempted at collection. These samples were collected for previous population genetic analyses or as controls in disease association studies[27,29,30,33,43,44]. Ethical approval for the present study was provided by the UCL/UCLH ethics committee (UK) as well as by ethics committees in the countries where the samples were collected. Most analyses were carried out using a dataset that also included genotype information for 160 Europeans, 123 Africans and 463 Native Americans (from 26 Amerindian populations, samples size 7–25). Data for the European, African and five of the Native American populations are from the HGDP-CEPH human genome diversity panel database (v 1.0) (http://www.cephb.fr/hgdp-cephdb/)[22]. Data for the 21 additional Native American populations are from Wang et al.[23]. The approximate location of the Mestizo and Native American populations included in the analyses is shown in Figure 1. Additional geographic and demographic information for the urban areas sampled is shown in Table S1. Sampling sites mostly correspond to one location. When more than one location was sampled in a given region, the census information provided is the sum of these locations and the sample was given a regional denomination (e.g. the Rio Grande do Sul (RGS) sample was collected in the cities of Bagé and Alegrete, in the Brazilian state of RGS).

Marker Data

Individual genotype data were collected by the Marshfield Foundation Mammalian Genotyping Service (http://research.marshfieldclinic.org/genetics/) for 751 microsatellites distributed across all 22 autosomes and 35 markers on the X-chromosome.

The markers examined are included in Marshfield Screening Sets 16 and 54, commonly employed in linkage studies. Standardization of allele calls for compilation of datafiles combining genotypes for 678 autosomal and 29 X-chromosome markers in Mestizos, Europeans, Africans and Native Americans was performed as detailed in Wang et. al. [23]. For X-chromosome data, males were treated as diploids with one missing allele at each locus.

Admixture Analysis

Estimation of individual ancestry proportions was performed with the programs STRUCTURE[25,45] and ADMIXMAP[24]. Since very similar estimates were obtained with both programs, only those obtained with STRUCTURE are reported here. Replicate runs of STRUCTURE used a burn-in period of 20,000 iterations followed by an additional 10,000 iterations from which parameter estimates were obtained. Ten replicate runs were carried out and the average parameter estimate retained. Population admixture proportions and mean time since admixture were calculated from the individual estimates. Spearman rank correlations (two-sided) and Wilcoxon signed rank tests between ancestry estimates and other population parameters were applied using the R statistical package (http://www.r-project.org). STRUCTURE runs used an admixture model with correlated allele frequencies and with individuals from ancestral populations assigned to K predetermined clusters (so-called "supervised analysis"). K was varied in order to examine different groupings of Native American populations while considering Europeans and Africans as single independent clusters: $K = 3$ when grouping all Native American data into a single cluster, $K = 7$ when Native American populations were grouped into five linguistic stocks and $K = 28$ when each Native American population was considered independently. Population assignment to linguistic stocks followed the linguistic classification proposed by Ruhlen (1991)[40].

Population Diversity Estimates

For each population, heterozygosity was computed for each locus using the unbiased estimator of Weir (1996)[46], and the average across loci was taken as the population estimate. Calculation of F_{ST} was performed using eq. 5.3 of Weir (1996)[46]. The expected heterozygosity (I) for the Mestizo was calculated using the expression of Rosenberg and Huang (personal communication):

$$I = \gamma^2 I_A + (1-\gamma)^2 I_B + \gamma(1-\gamma)(I_A + I_B)\frac{1+F}{1-F}$$

Where I_A and I_B are the observed heterozygosities of European and Native American populations, F the F_{ST} estimated between Europeans and Africans, and γ the proportion of European ancestry in the Mestizo. A related expression for the expected heterozygosity after admixture of three populations (i.e. including Africa) did not produce a significantly better fit with the data analysed here.

Native American Ancestry and Geographic Distance between Populations

For each Mestizo population, we computed a Pearson correlation coefficient between the Native American ancestry components (as shown in Figure 6) and the logarithm of the distance to the corresponding Native population (using the population coordinates shown in Table S1 and those reported in Wang et al.[23]). Significance of correlations was evaluated using the standard one-sided t-distribution transformation. A mean weighted correlation coefficient was obtained by averaging correlations over mestizo populations after weighting for sample size. Besides Euclidean distances, we computed effective distances using PATHMATRIX [47] and employing a 1:10 coastal/inland cost ratio (i.e. therefore assuming that coastlines facilitated migration) (see Wang et al. (2007)[23]for details and rationale).

The genotypes analyzed here are included in Dataset S1, available online.

Supporting Information

Figure S1 Distribution of admixture estimates for individuals from 11 Mestizo populations. The position of each blue dot on the triangle plot indicates the proportion of European, Native American and African ancestry estimated for each individual in the population.
Found at: doi:10.1371/journal.pgen.1000037.s001 (0.34 MB XLS)

Table S1 Location of sampling site, current population size and relative pre-Columbian population density in the region. * Representing about 10, 5 and 1 individuals per square mile, according to available estimates[13;14]. Population size information from compilation in(http://www.citypopulation.de/cities.html).
Found at: doi:10.1371/journal.pgen.1000037.s002 (0.03 MB XLS)

Table S2 Mean ancestry proportions, variance (var.) and standard errors (s.e.) for individuals from thirteen Mestizo populations.
Found at: doi:10.1371/journal.pgen.1000037.s003 (0.03 MB XLS)

Table S3 Mean number of generations (s.d.) to admixture in 13 mestizo populations based on 678 autosomal microsatellite markers.
Found at: doi:10.1371/journal.pgen.1000037.s004 (0.03 MB XLS)

Table S4 Mean ancestry proportions, variance (var.) and standard errors (s.e.) for individuals from thirteen Mestizo populations estimated with Native Americans subdivided according to linguistic affiliation.
Found at: doi:10.1371/journal.pgen.1000037.s005 (0.03 MB XLS)

Table S5 Mean ancestry proportions, variance (var.) and standard errors (s.e.) for individuals from thirteen Mestizo populations estimated with each Native American population considered individually.
Found at: doi:10.1371/journal.pgen.1000037.s006 (0.05 MB XLS)

Dataset S1 Genotype data used in Geographic patterns of genome admixture in Latin American Mestizos.
Found at: doi:10.1371/journal.pgen.1000037.s007 (1.51 MB)

Acknowledgments

Author Contributions

Conceived and designed the experiments: ARL. Analyzed the data: SW NR LE ARL. Contributed reagents/materials/analysis tools: WR MP GB

CG GP GM KH AH BC HN WK RB JM NF MB FS MP LT JD EA GB NB EL FR ARL. Wrote the paper: ARL. Designed and produced the figures SW NR ARL. Supervised data analysis: LE ARL.

Reference List

1. Zhu X, Cooper RS, Elston RC (2004) Linkage analysis of a complex disease through use of admixed populations. Am J Hum Genet 74: 1136–1153.
2. Smith MW, O'Brien SJ (2005) Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. Nat Rev Genet 6: 623–632.
3. McKeigue PM (2005) Prospects for admixture mapping of complex traits. Am J Hum Genet 76: 1–7.
4. Tang H, Coram M, Wang P, Zhu X, Risch N (2006) Reconstructing genetic ancestry blocks in admixed individuals. Am J Hum Genet 79: 1–12.
5. Seldin MF (2007) Admixture mapping as a tool in gene discovery. Curr Opin Genet Dev 17: 177–181.
6. Zhu X, Luke A, Cooper RS, Quertermous T, Hanis C, et al. (2005) Admixture mapping for hypertension loci with genome-scan markers. Nat Genet 37: 177–181.
7. Smith MW, Patterson N, Lautenberger JA, Truelove AL, McDonald GJ, et al. (2004) A high-density admixture map for disease gene discovery in african americans. Am J Hum Genet 74: 1001–1013.
8. Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, et al. (2006) A genomewide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. Am J Hum Genet 79: 640–649.
9. Reich D, Patterson N, De Jager PL, McDonald GJ, Waliszewska A, et al. (2005) A whole-genome admixture scan finds a candidate locus for multiple sclerosis susceptibility. Nat Genet 37: 1113–1118.
10. Deo RC, Patterson N, Tandon A, McDonald GJ, Haiman CA, et al. (2007) A High-Density Admixture Scan in 1,670 African Americans with Hypertension. PLoS Genet 3: e196.
11. Reich D, Patterson N, Ramesh V, De Jager PL, McDonald GJ, et al. (2007) Admixture mapping of an allele affecting interleukin 6 soluble receptor and interleukin 6 levels. Am J Hum Genet 80: 716–726.
12. Gonzalez BE, Borrell LN, Choudhry S, Naqvi M, Tsai HJ, et al. (2005) Latino populations: a unique opportunity for the study of race, genetics, and social environment in epidemiological research. Am J Public Health 95: 2161–2168.
13. Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, et al. (2007) A genomewide admixture mapping panel for hispanic/latino populations. Am J Hum Genet 80: 1171–1178.
14. Price AL, Patterson N, Yu F, Cox DR, Waliszewska A, et al. (2007) A genomewide admixture map for latino populations. Am J Hum Genet 80: 1024–1036.
15. Tian C, Hinds DA, Shigeta R, Adler SG, Lee A, et al. (2007) A genomewide single-nucleotide-polymorphism panel for Mexican American admixture mapping. Am J Hum Genet 80: 1014–1023.
16. Morner M (1967) Race Mixture in the History of Latin America. Little Brown & Company.
17. Sanchez-Albornoz N (1974) The population of Latin America: a history. Berkeley: University of California Press.
18. Burkholder MA, Johnson LL (2003) Colonial Latin America. Oxford University Press.
19. Collier S, Skidmore TE, Blakemore H (1992) Cambridge Encyclopedia of Latin America and the Caribbean. Cambridge: Cambridge University Press.
20. Sans M (2000) Admixture studies in Latin America: from the 20th to the 21st century. Hum Biol 72: 155–177.
21. Salzano FM, Bortolini MC (2002) The Evolution and Genetics of Latin American Populations. Cambridge University Press.
22. Cann HM, de Toma C, Cazes L, Legrand MF, Morel V, et al. (2002) A human genome diversity cell line panel. Science 296: 261–262.
23. Wang S, Lewis CM, Jakobsson M, Ramachandran S, Ray N, et al. (2007) Genetic Variation and Population Structure in Native Americans. PLoS Genet 3: e185.
24. Hoggart CJ, Shriver MD, Kittles RA, Clayton DG, McKeigue PM (2004) Design and analysis of admixture mapping studies. Am J Hum Genet 74: 965–978.
25. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587.
26. Freeman AR, Hoggart CJ, Hanotte O, Bradley DG (2006) Assessing the relative ages of admixture in the bovine hybrid zones of Africa and the Near East using X chromosome haplotype mosaicism. Genetics 173: 1503–1510.
27. Escamilla MA, Spesny M, Reus VI, Gallegos A, Meza L, et al. (1996) Use of linkage disequilibrium approaches to map genes for bipolar disorder in the Costa Rican population. Am J Med Genet 67: 244–253.
28. Service S, Deyoung J, Karayiorgou M, Roos JL, Pretorious H, et al. (2006) Magnitude and distribution of linkage disequilibrium in population isolates and implications for genome-wide association studies. Nat Genet 38: 556–560.
29. Alfaro EL, Dipierri JE, Gutierrez NI, Vullo CM (2005) Genetic structure and admixture in urban populations of the Argentine North-West. Ann Hum Biol 32: 724–737.
30. Marrero AR, Bravi C, Stuart S, Long JC, Pereira das Neves LF, et al. (2007) Pre- and post-Columbian gene and cultural continuity: the case of the Gaucho from southern Brazil. Hum Hered 64: 160–171.
31. Bonilla C, Parra EJ, Pfaff CL, Dios S, Marshall JA, et al. (2004) Admixture in the Hispanics of the San Luis Valley, Colorado, and its implications for complex trait gene mapping. Ann Hum Genet 68: 139–153.
32. Seldin MF, Tian C, Shigeta R, Scherbarth HR, Silva G, et al. (2007) Argentine population genetic structure: large variance in Amerindian contribution. Am J Phys Anthropol 132: 455–462.
33. Carvajal-Carmona LG, Soto ID, Pineda N, Ortiz-Barrientos D, Duque C, et al. (2000) Strong Amerind/White Sex Bias and a Possible Sephardic Contribution among the Founders of a Population in Northwest Colombia. Am J Hum Genet 67: 1287–1295.
34. Green LD, Derr JN, Knight A (2000) mtDNA affinities of the peoples of North-Central Mexico. Am J Hum Genet 66: 989–998.
35. Carvajal-Carmona LG, Ophoff R, Service S, Hartiala J, Molina J, et al. (2003) Genetic demography of Antioquia (Colombia) and the Central Valley of Costa Rica. Hum Genet 112: 534–541.
36. Bedoya G, Montoya P, Garcia J, Soto I, Bourgeois S, et al. (2006) Admixture dynamics in Hispanics: a shift in the nuclear genetic ancestry of a South American population isolate. Proc Natl Acad Sci U S A 103: 7234–7239.
37. Dipierri JE, Alfaro E, Martinez-Marignac VL, Bailliet G, Bravi CM, et al. (1998) Paternal directional mating in two Amerindian subpopulations located at different altitudes in northwestern Argentina. Hum Biol 70: 1001–1010.
38. Sans M, Weimer TA, Franco MH, Salzano FM, Bentancor N, et al. (2002) Unequal contributions of male and female gene pools from parental populations in the African descendants of the city of Melo, Uruguay. Am J Phys Anthropol 118: 33–44.
39. Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, et al. (1998) Estimating African American admixture proportions by use of population- specific alleles. Am J Hum Genet 63: 1839–1851.
40. Ruhlen M (1991) A Guide to the World's Languages. Stanford, Calif.: Stanford University Press.
41. Villa M, Rodriguez J (1996) Demographic trends in Latin America's metropolises, 1950–1990. In: Gilbert A, ed. The Mega-city in Latin America. Tokyo: United Nations University Press. pp 25–52.
42. Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, et al. (2004) Methods for high-density admixture mapping of disease genes. Am J Hum Genet 74: 979–1000.
43. Henriquez H, Moraga M, Llop E, Rothhammer F (2004) [Molecular and genetic characterization of Changos descendants living in Paposo Cove]. Rev Med Chil 132: 663–672.
44. Campos-Sanchez R, Barrantes R, Silva S, Escamilla M, Ontiveros A, et al. (2006) Genetic structure analysis of three Hispanic populations from Costa Rica, Mexico, and the southwestern United States using Y-chromosome STR markers and mtDNA sequences. Hum Biol 78: 551–563.
45. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959.
46. Weir BS (1996) Genetic Data Analysis II. Sunderland, MA: Sinauer.
47. Ray N (2005) PATHMATRIX: a geographical information system tool to compute effective distances among samples. Molecular Ecology Notes 5: 177–180.