

Geographic patterns of genome admixture in Latin American Mestizos

Sijia Wang¹, Nicolas Ray², Winston Rojas³, Maria V. Parra³, Gabriel Bedoya³, Carla Gallo⁴, Giovanni Poletti⁴, Guido Mazzotti⁵, Kim Hill⁶, Ana M. Hurtado⁶, Beatriz Camrena⁷, Humberto Nicolini⁷, William Klitz⁸, Ramiro Barrantes⁹, Julio A. Molina¹⁰, Nelson Freimer¹⁰, Maria Cátira Bortolini¹¹, Francisco M. Salzano¹¹, Maria L. Petzl-Erler¹², Luiza T. Tsuneto¹², José E. Dipierri¹³, Emma L. Alfaro¹³, Graciela Bailliet¹⁴, Nestor O. Bianchi¹⁴, Elena Llop¹⁵, Francisco Rothhammer^{15,16}, Laurent Excoffier², and Andrés Ruiz-Linares¹

¹ The Galton Laboratory, Department of Biology, University College London, 4 Stephenson Way, London NW1 2HE, 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, USA.

⁷ Departamento de Ciencias Genómicas, Universidad Autónoma de la Ciudad de México. México D.F., México.

⁸ University of California, School of Public Health, Berkeley and Public Health Institute, Oakland, USA.

⁹ Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica.

¹⁰ Center for Neurobehavioral Genetics, University of California, Los Angeles, USA.

¹¹ Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil.

¹² Departamento de Genética, Universidade Federal do Paraná, Curitiba, Brazil

¹³ Instituto de Biología de la Altura, Facultad de Humanidades y Ciencias Sociales, Universidad Nacional de Jujuy, San Salvador de Jujuy, Argentina.

¹⁴ Laboratory of Human Molecular Population Genetics, IMBICE, La Plata, Argentina.

¹⁵ Programa de Genética Humana, Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile, Santiago, Chile.

¹⁶ Instituto de Alta Investigación, Universidad de Tarapacá, Arica, Chile.

Corresponding author: Andrés Ruiz-Linares. Department of Biology. University College London. 4 Stephenson Way. London NW1 2HE. United Kingdom. Tel: +44 (0)20 7679 5049. Fax: +44 (0)20 7679 5052. Email: a.ruizlin@ucl.ac.uk

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.

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 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 Supplementary TableS1). 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 and 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 2A (the corresponding values and standard errors of these estimates are presented in Supplementary TableS2). 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 (Supplementary FigureS1). 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 (Supplementary TableS1), 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 (Supplementary TableS3), 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 2A). 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 2B). 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} ($\rho=0.786$ $P<0.01$; Figure 2B and Supporting online material). 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 2C, Supplementary FigureS1 and TableS2), 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 2A, 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 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 2A, 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 2A 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 2A).

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 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 therefore investigated whether it is possible to detect such an underlying genetic differentiation amongst Mestizos by taking advantage of the fact that we sampled a wide range of Mestizo populations and also that data for Native American populations from the same areas is available[23]. We therefore carried out 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 3A 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 Supplementary TableS4). 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 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 3B, Figure 1 and Supplementary TableS5). 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 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 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 3B). These correlations are negative for all Mestizo populations (Figure 3C), 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 the differentiation of the Native American component in Mestizos 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 (Supplementary TableS1), both in the linguistic-based analysis (Figure 3A, $\rho=-0.611$, $P<0.05$) and in the population-based analysis (Figure 3B, $\rho=-0.661$ $P<0.05$). 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 recent 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 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 2C). These populations could be especially 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 considerable variation in rates of admixture between populations, and sometimes also within populations, needs to be taken into account when designing admixture mapping studies in specific Latin American populations. Nevertheless, we anticipate that admixture mapping in Mestizos will prove a very valuable strategy for elucidating the genetic basis of various 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 (manuscript in preparation):

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

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 3B) and the logarithm of the distance to the corresponding Native population (using the population coordinates shown in Table 1 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).

Acknowledgments

We thank Cheryl Winkler, Phabiola Herrera and David E. Ruiz R. for assistance with various aspects of this work. We are also grateful to Rolando Gonzalez-José, Marc Feldman and Noah Rosenberg for comments on the manuscript and to Daniel Falush and Mattias Jakobsson for assistance with STRUCTURE 2.0. Funding for this work was provided by the following sources: A.R-L (NIH R01NS043538); F.R (Fondecyt Proyecto 1050595); M.C.B., M.L.P.E. and F.M.S. (Institutos do Milênio and Apoio a Núcleos de Excelência Programs, Conselho Nacional de Desenvolvimento Científico e Tecnológico). M.C.B. and F.M.S. (Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul). M.L.P.E. (Fundação Araucária). N.B.F. (NIH R01MH049499, R01NS37484), J.E.D., E.L.A., N.O.B. and G.Bailliet (CONICET, CICPBA and ANPCyT Argentina). G.Bedoya (Colciencias and Universidad de Antioquia). N.R and L.E. (Swiss NSF grant No. 3100-112072). S.W. was supported by an Overseas Research Studentship (UK) and a K.C. Wong Studentship. Genotyping at the Marshfield Foundation Mammalian Genotyping Service was supported by the National Heart, Lung, and Blood Institute. The data analyzed here is available upon request.

Author contributions

ARL designed the study. ARL, GBe WR, MVP, JM, NF, CG, GP, WK, KH, AMH, BC, HN, RB, MCB, FMS, MLPE, LTT, JED, ELA, GBa, NOB, EL and FR contributed DNA samples, which were prepared for genotyping by SW. LE and ARL supervised data

analysis. SW, NR and ARL analyzed the data. SW, NR and ARL designed and produced the figures. ARL wrote the text of the paper and incorporated contributions from all co-authors.

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, Mosley T, Gu CC, Tang H, Rao DC, Risch N, Weder A (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, 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: 1001-1013.
8. Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, Seldin MF (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, Tandon A, Lincoln RR, DeLoa C, Fruhan SA, Cabre P, Bera O, Semana G, Kelly MA, Francis DA, Ardlie K, Khan O, Cree BA, Hauser SL, Oksenberg JR, Hafler DA (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, Ardlie K, Henderson BE, Henderson SO, Reich D (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, Tandon A, Choy E, Hu D, Tamraz B, Pawlikowska L, Wassel-Fyr C, Huntsman S, Waliszewska A, Rossin E, Li R, Garcia M, Reiner A, Ferrell R, Cummings S, Kwok PY,

- Harris T, Zmuda JM, Ziv E (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, Rodriguez-Santana JR, Chapela R, Rogers SD, Mei R, Rodriguez-Cintrón W, Arena JF, Kittles R, Perez-Stable EJ, Ziv E, Risch N (2005) Latino populations: a unique opportunity for the study of race, genetics, and social environment in epidemiological research. *Am J Public Health* 95: 2161-2168.
 13. Mao X, Bigham AW, Mei R, Gutierrez G, Weiss KM, Brutsaert TD, Leon-Velarde F, Moore LG, Vargas E, McKeigue PM, Shriver MD, Parra EJ (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, McDonald GJ, Tandon A, Schirmer C, Neubauer J, Bedoya G, Duque C, Villegas A, Bortolini MC, Salzano FM, Gallo C, Mazzotti G, Tello-Ruiz M, Riba L, Aguilar-Salinas CA, Canizales-Quinteros S, Menjivar M, Klitz W, Henderson B, Haiman CA, Winkler C, Tusie-Luna T, Ruiz-Linares A, Reich D (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, Pahl MV, Silva G, Belmont JW, Hanson RL, Knowler WC, Gregersen PK, Ballinger DG, Seldin MF (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, Piouffre L, Bodmer J, Bodmer WF, Bonne-Tamir B, Cambon-Thomsen A, Chen Z, Chu J,

- Carcassi C, Contu L, Du R, Excoffier L, Ferrara GB, Friedlaender JS, Groot H, Gurwitz D, Jenkins T, Herrera RJ, Huang X, Kidd J, Kidd KK, Langaney A, Lin AA, Mehdi SQ, Parham P, Piazza A, Pistillo MP, Qian Y, Shu Q, Xu J, Zhu S, Weber JL, Greely HT, Feldman MW, Thomas G, Dausset J, Cavalli-Sforza LL (2002) A human genome diversity cell line panel. *Science* 296: 261-262.
23. Wang S, Lewis CM, Jakobsson M, ramachandran s, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A (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, Molina J, Sandkuijl LA, Fournier E, Leon PE, Smith LB, Freimer NB (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, Pretorius H, Bedoya G, Ospina J, Ruiz-Linares A, Macedo A, Palha JA, Heutink P, Aulchenko Y, Oostra B, van Duijn C, Jarvelin MR, Varilo T, Peddle L, Rahman P, Piras G, Monne M, Murray S, Galver L, Peltonen L, Sabatti C, Collins A, Freimer N (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, Kommers T, Carvalho CM, Pena SD, Ruiz-Linares A, Salzano FM, Catira BM (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, Hamman RF, Ferrell RE, Hoggart CL, McKeigue PM, Shriver MD (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, Belmont JW, Kittles R, Gamron S, Allevi A, Palatnik SA, Alvarellos A, Paira S, Caprarulo C, Guilleron C, Catoggio LJ, Prigione C, Berbotto GA, Garcia MA, Perandones CE, Pons-Estel BA, Alarcon-Riquelme ME (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, Ospina-Duque J, McCarthy M, Montoya P, Alvarez VM, Bedoya G, Ruiz-Linares A (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, Leon P, Ospina J, Bedoya G, Freimer N, Ruiz-Linares A (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., Carvajal L, Labuda D, Alvarez VM, Ospina J, Hedrick PW, Ruiz-Linares A (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, Cejas S, Bianchi NO (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, Alvarez I, Bianchi NO, Chakraborty R (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, Cooper R, Forrester T, Allison DB, Deka R, Ferrell RE, Shriver MD (1998) Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet* 63: 1839-1851.
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, editors. *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, 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: 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, Nicolini H, Mendoza R, Munoz R, Raventos H (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*. Sinauer, Sunderland, MA.
47. Ray N (2005) PATHMATRIX: a geographical information system tool to compute effective distances among samples. *Molecular Ecology Notes* 5: 177-180.

FIGURE LEGENDS

Figure 1. Approximate geographic location of the Mestizo populations examined here and of the Native American populations also 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.

Figure 2. (A) 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 Supplementary TableS2. (B) Heterozygosity vs. proportion of European ancestry in Mestizo populations. The heterozygosity predicted from the estimated ancestry of a population was calculated as described in the Supporting online material. (C) Distribution of admixture estimates for individuals from Mexico City and Quetalmahue. The position of each green 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 Supplementary online information (FigureS1). In (A) to (C) 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. (in press)[23].

Figure 3. Regional Native American ancestry of 13 Mestizo populations. 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). Panel (A): when considering the major Native American linguistic stocks present in the region (except for Ge-Pano-Carib which is not included as it is represented here only by one population; the Kaingang). Panel (B): when considering each individual Native population; these are coded in shades of a color corresponding to the main linguistic stocks shown in panel (A) and Figure 1. Panel (C) shows for each Mestizo population the Pearson correlation coefficients 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 panel B). Correlation coefficients with associated *P*-values <0.05 are shown as filled symbols.

Fig 1

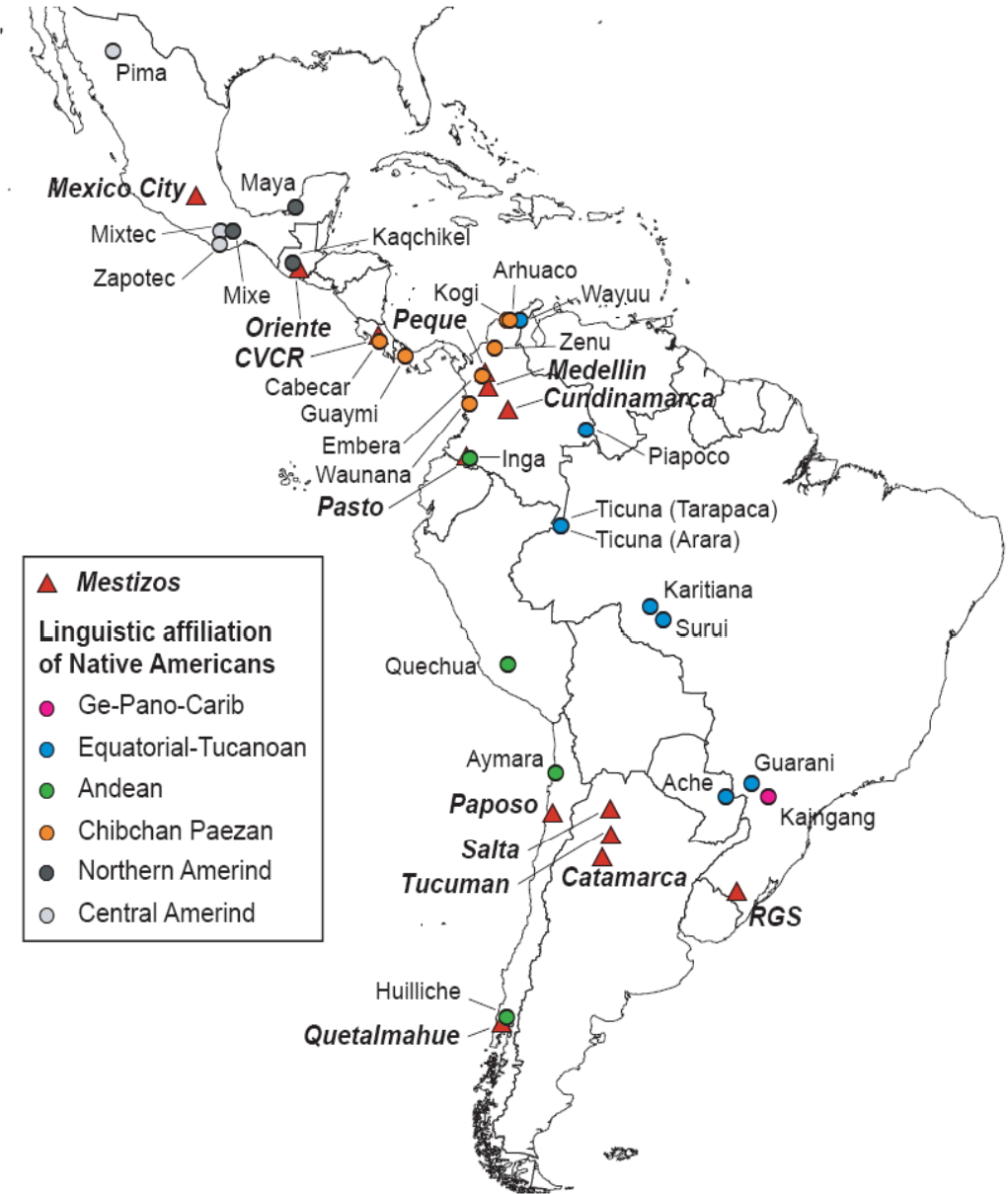


Fig 2A

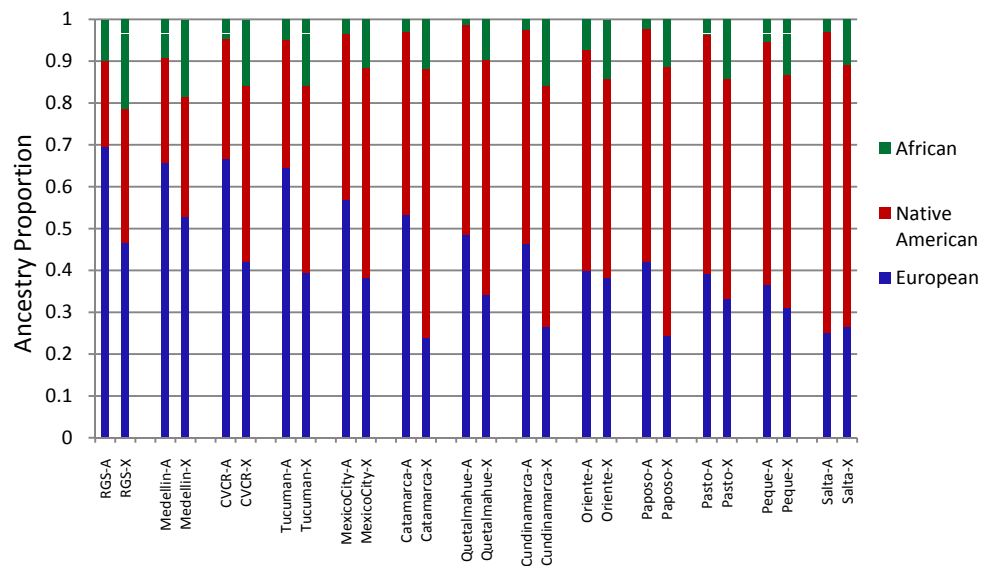


Fig 2B

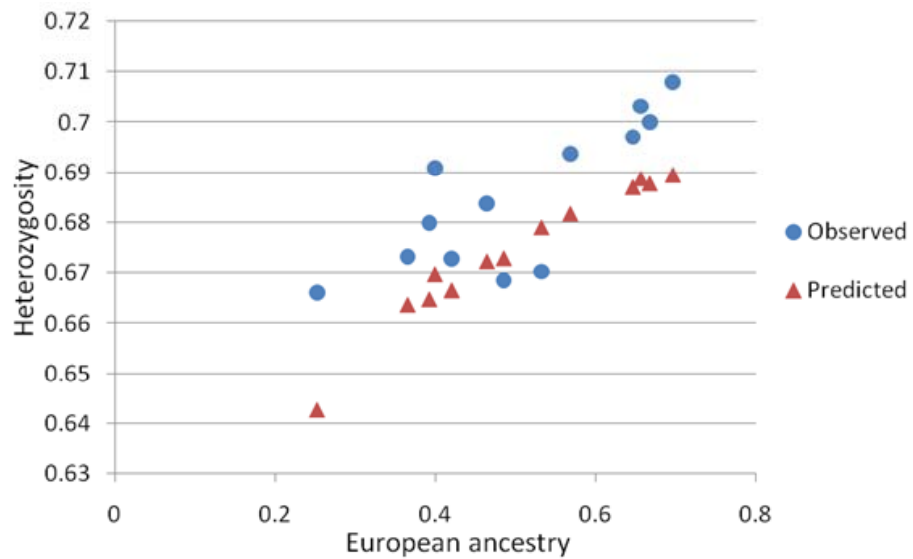


Fig 2C

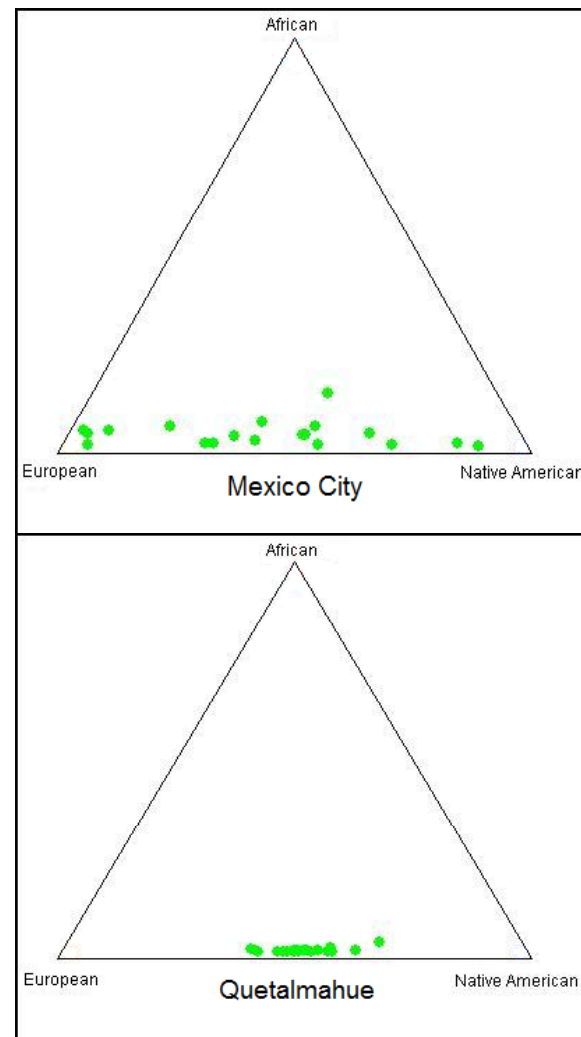


Fig 3A

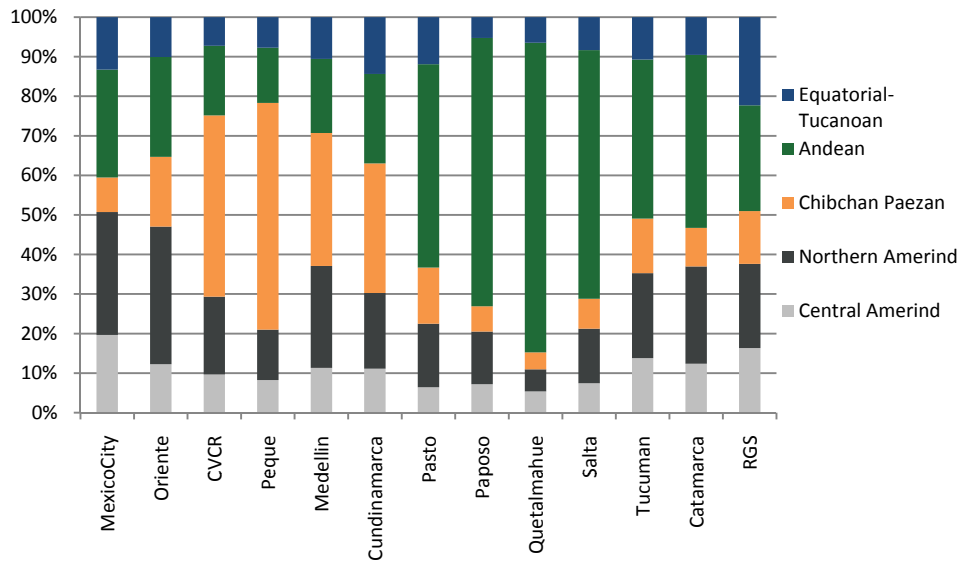


Fig 3B

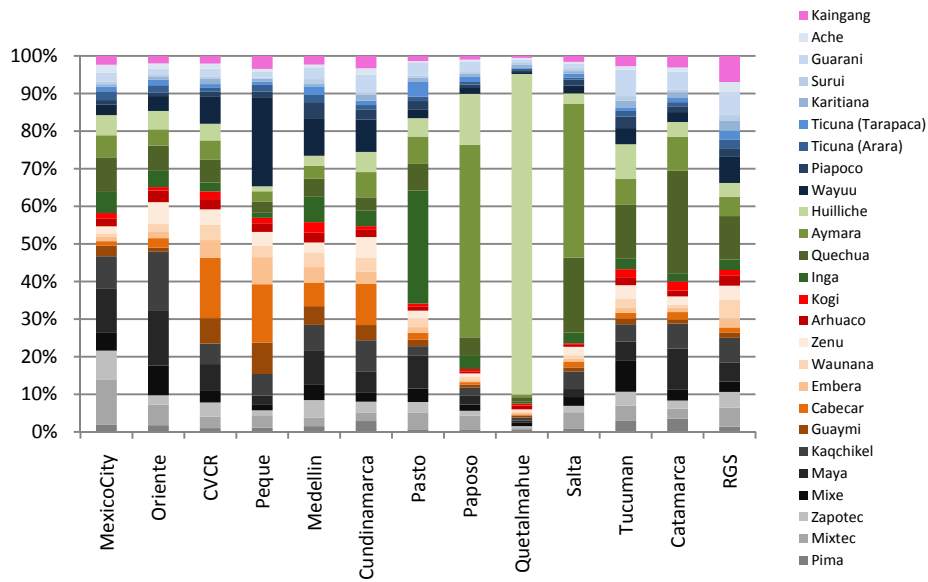
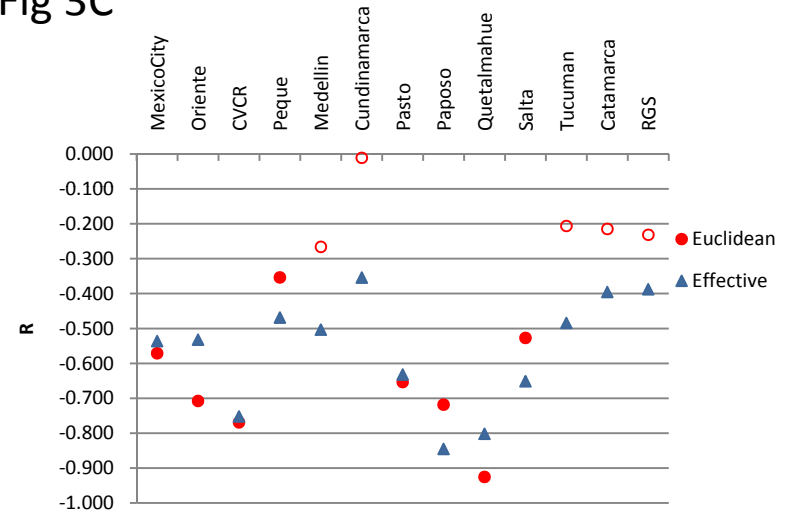


Fig 3C



SUPPORTING ONLINE MATERIAL

TableS1 Location of sampling site, current population size and relative pre-Columbian population density in the region.

| Sampling site | (n) | Country | Latitude | Longitude | Population | Pre-Columbian density* |
|---------------|-----|------------|----------|-----------|------------|------------------------|
| Mexico City | 19 | Mexico | 19°24'N | 99°12'W | 18,000,000 | High |
| Oriente | 20 | Guatemala | 14°38'N | 90°31'W | 108,000 | High |
| CVCR | 20 | Costa Rica | 09°56'N | 84°05'W | 2,500,000 | Medium |
| Peque | 20 | Colombia | 07°19'N | 75°50'W | 2000 | Medium |
| Medellin | 20 | Colombia | 06°15'N | 75°35'W | 2,500,000 | Medium |
| Cundinamarca | 19 | Colombia | 04°36'N | 74°05'W | 2,500,000 | Medium |
| Pasto | 19 | Colombia | 01°13'N | 77°17'W | 365,000 | High |
| Paposo | 20 | Chile | 25°03'S | 70°16'W | 1000 | High |
| Quetalmahue | 20 | Chile | 41°28'S | 73°32'W | 1000 | High |
| Salta | 19 | Argentina | 24°47'S | 65°25'W | 469,000 | High |
| Tucuman | 19 | Argentina | 26°49'S | 65°13'W | 738,000 | Medium |
| Catamarca | 14 | Argentina | 28°28'S | 65°47'W | 171,000 | Medium |
| RGS | 20 | Brazil | 31°00'S | 54°00'W | 200,000 | Low |

* Representing about 10, 5 and 1 individuals per square mile, according to available estimates^{13;14}.

TableS2 Mean ancestry proportions, variance (var.) and standard errors (s.e.) for individuals from thirteen Mestizo populations.

| Population | AUTOSOMES | | | | | | | | | X-CHROMOSOME | | | | | | | | |
|--------------|-----------|-------|-------|-----------------|-------|-------|---------|-------|-------|--------------|-------|-------|-----------------|-------|-------|---------|-------|-------|
| | European | | | Native American | | | African | | | European | | | Native American | | | African | | |
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.568 | 0.071 | 0.061 | 0.398 | 0.072 | 0.061 | 0.034 | 0.001 | 0.006 | 0.383 | 0.068 | 0.060 | 0.501 | 0.076 | 0.063 | 0.116 | 0.003 | 0.013 |
| Oriente | 0.399 | 0.041 | 0.045 | 0.528 | 0.047 | 0.049 | 0.073 | 0.003 | 0.011 | 0.383 | 0.056 | 0.053 | 0.474 | 0.065 | 0.057 | 0.142 | 0.008 | 0.020 |
| CVCR | 0.667 | 0.011 | 0.024 | 0.287 | 0.007 | 0.019 | 0.046 | 0.001 | 0.008 | 0.420 | 0.038 | 0.043 | 0.420 | 0.028 | 0.038 | 0.159 | 0.008 | 0.020 |
| Peque | 0.365 | 0.021 | 0.032 | 0.582 | 0.022 | 0.033 | 0.053 | 0.001 | 0.008 | 0.311 | 0.045 | 0.047 | 0.557 | 0.070 | 0.059 | 0.133 | 0.009 | 0.021 |
| Medellin | 0.656 | 0.019 | 0.031 | 0.253 | 0.006 | 0.018 | 0.091 | 0.006 | 0.017 | 0.529 | 0.048 | 0.049 | 0.285 | 0.028 | 0.037 | 0.185 | 0.022 | 0.033 |
| Cundinamarca | 0.464 | 0.018 | 0.031 | 0.511 | 0.019 | 0.031 | 0.025 | 0.000 | 0.004 | 0.264 | 0.025 | 0.036 | 0.577 | 0.033 | 0.042 | 0.159 | 0.015 | 0.028 |
| Pasto | 0.392 | 0.025 | 0.036 | 0.572 | 0.027 | 0.038 | 0.036 | 0.003 | 0.013 | 0.331 | 0.047 | 0.050 | 0.526 | 0.052 | 0.052 | 0.143 | 0.010 | 0.023 |
| Paposo | 0.420 | 0.006 | 0.018 | 0.557 | 0.008 | 0.020 | 0.023 | 0.000 | 0.004 | 0.244 | 0.017 | 0.030 | 0.642 | 0.014 | 0.027 | 0.114 | 0.003 | 0.012 |
| Quetalmahue | 0.485 | 0.005 | 0.017 | 0.502 | 0.005 | 0.015 | 0.013 | 0.000 | 0.002 | 0.342 | 0.029 | 0.038 | 0.561 | 0.037 | 0.043 | 0.098 | 0.003 | 0.012 |
| Salta | 0.252 | 0.033 | 0.042 | 0.719 | 0.036 | 0.044 | 0.030 | 0.001 | 0.005 | 0.264 | 0.041 | 0.047 | 0.627 | 0.047 | 0.049 | 0.109 | 0.002 | 0.010 |
| Tucuman | 0.646 | 0.037 | 0.044 | 0.305 | 0.035 | 0.043 | 0.049 | 0.001 | 0.008 | 0.395 | 0.055 | 0.054 | 0.447 | 0.063 | 0.057 | 0.158 | 0.017 | 0.030 |
| Catamarca | 0.532 | 0.007 | 0.022 | 0.437 | 0.007 | 0.022 | 0.030 | 0.000 | 0.006 | 0.238 | 0.029 | 0.046 | 0.643 | 0.034 | 0.050 | 0.119 | 0.002 | 0.012 |
| RGS | 0.696 | 0.024 | 0.034 | 0.204 | 0.008 | 0.020 | 0.099 | 0.014 | 0.026 | 0.466 | 0.046 | 0.048 | 0.321 | 0.036 | 0.043 | 0.213 | 0.017 | 0.029 |

Table S3 Mean number of generations (s.d.) to admixture in 13 mestizo populations based on 678 autosomal microsatellite markers.

| Population | mean | (s.d.) |
|-------------------|-------------|---------------|
| Mexico City | 9.04 | 1.44 |
| Oriente | 8.95 | 1.15 |
| CVCR | 14.42 | 2.32 |
| Peque | 7.54 | 0.86 |
| Medellin | 9.65 | 1.38 |
| Cundinamarca | 12.31 | 1.56 |
| Pasto | 8.88 | 1.08 |
| Paposo | 8.37 | 1.00 |
| Quetalmahue | 9.39 | 1.21 |
| Salta | 6.21 | 0.93 |
| Tucuman | 10.09 | 1.56 |
| Catamarca | 11.00 | 2.03 |
| RGS | 6.57 | 0.86 |

TableS4 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.

| Population | European | | | African | | | Northern Amerind | | | Central Amerind | | | Chibchan Paezan | | | Andean | | | Equatorial-Tucanoan | | |
|--------------|----------|-------|-------|---------|-------|-------|------------------|-------|-------|-----------------|-------|-------|-----------------|-------|-------|--------|-------|-------|---------------------|-------|-------|
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.546 | 0.073 | 0.062 | 0.030 | 0.001 | 0.007 | 0.131 | 0.011 | 0.024 | 0.083 | 0.006 | 0.017 | 0.037 | 0.001 | 0.008 | 0.115 | 0.008 | 0.020 | 0.056 | 0.004 | 0.015 |
| Oriente | 0.383 | 0.042 | 0.046 | 0.063 | 0.002 | 0.011 | 0.193 | 0.013 | 0.025 | 0.068 | 0.003 | 0.012 | 0.098 | 0.009 | 0.021 | 0.140 | 0.011 | 0.023 | 0.056 | 0.003 | 0.012 |
| CVCR | 0.664 | 0.012 | 0.024 | 0.047 | 0.001 | 0.009 | 0.057 | 0.001 | 0.008 | 0.028 | 0.000 | 0.004 | 0.133 | 0.005 | 0.016 | 0.051 | 0.002 | 0.009 | 0.021 | 0.000 | 0.002 |
| Peque | 0.364 | 0.022 | 0.033 | 0.042 | 0.001 | 0.008 | 0.076 | 0.003 | 0.012 | 0.049 | 0.002 | 0.009 | 0.341 | 0.012 | 0.025 | 0.083 | 0.004 | 0.014 | 0.046 | 0.001 | 0.006 |
| Medellin | 0.655 | 0.019 | 0.031 | 0.089 | 0.006 | 0.018 | 0.066 | 0.005 | 0.015 | 0.029 | 0.001 | 0.006 | 0.086 | 0.003 | 0.012 | 0.048 | 0.002 | 0.009 | 0.027 | 0.000 | 0.004 |
| Cundinamarca | 0.461 | 0.019 | 0.032 | 0.018 | 0.000 | 0.004 | 0.100 | 0.005 | 0.016 | 0.058 | 0.002 | 0.010 | 0.171 | 0.009 | 0.022 | 0.118 | 0.006 | 0.018 | 0.075 | 0.005 | 0.016 |
| Pasto | 0.367 | 0.029 | 0.039 | 0.028 | 0.004 | 0.014 | 0.097 | 0.009 | 0.022 | 0.039 | 0.001 | 0.008 | 0.086 | 0.004 | 0.014 | 0.311 | 0.021 | 0.033 | 0.072 | 0.002 | 0.011 |
| Paposo | 0.378 | 0.009 | 0.021 | 0.013 | 0.000 | 0.003 | 0.081 | 0.004 | 0.014 | 0.044 | 0.001 | 0.008 | 0.039 | 0.001 | 0.008 | 0.414 | 0.013 | 0.026 | 0.032 | 0.001 | 0.007 |
| Quetalmahue | 0.437 | 0.006 | 0.018 | 0.006 | 0.000 | 0.001 | 0.031 | 0.001 | 0.007 | 0.030 | 0.001 | 0.008 | 0.024 | 0.000 | 0.003 | 0.436 | 0.010 | 0.022 | 0.036 | 0.001 | 0.009 |
| Salta | 0.218 | 0.040 | 0.046 | 0.015 | 0.000 | 0.004 | 0.106 | 0.005 | 0.016 | 0.057 | 0.002 | 0.011 | 0.058 | 0.002 | 0.010 | 0.482 | 0.026 | 0.037 | 0.064 | 0.002 | 0.010 |
| Tucuman | 0.637 | 0.037 | 0.044 | 0.039 | 0.001 | 0.007 | 0.070 | 0.005 | 0.016 | 0.045 | 0.002 | 0.010 | 0.045 | 0.004 | 0.014 | 0.131 | 0.013 | 0.026 | 0.035 | 0.001 | 0.007 |
| Catamarca | 0.520 | 0.008 | 0.023 | 0.020 | 0.000 | 0.004 | 0.113 | 0.007 | 0.023 | 0.057 | 0.003 | 0.015 | 0.045 | 0.001 | 0.006 | 0.201 | 0.008 | 0.025 | 0.044 | 0.002 | 0.013 |
| RGS | 0.697 | 0.023 | 0.034 | 0.100 | 0.013 | 0.026 | 0.043 | 0.001 | 0.007 | 0.033 | 0.001 | 0.005 | 0.027 | 0.000 | 0.004 | 0.054 | 0.002 | 0.010 | 0.045 | 0.001 | 0.008 |

TableS5 Mean ancestry proportions, variance (var.) and standard errors (s.e.) for individuals from thirteen Mestizo populations estimated with each Native American population considered individually.

| Population | European | | | African | | | Pima | | | Mixtec | | | Zapotec | | | Mixe | | | Maya | | |
|--------------|-----------|-------|-------|---------|-------|-------|---------|-------|-------|--------|-------|-------|---------|-------|-------|-------|-------|-------|---------|-------|-------|
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.557 | 0.073 | 0.062 | 0.019 | 0.001 | 0.007 | 0.049 | 0.007 | 0.019 | 0.021 | 0.001 | 0.006 | 0.037 | 0.004 | 0.014 | 0.008 | 0.000 | 0.002 | 0.051 | 0.005 | 0.016 |
| Oriente | 0.399 | 0.046 | 0.048 | 0.047 | 0.002 | 0.011 | 0.081 | 0.014 | 0.027 | 0.044 | 0.006 | 0.017 | 0.086 | 0.018 | 0.030 | 0.010 | 0.000 | 0.002 | 0.030 | 0.005 | 0.015 |
| CVCR | 0.675 | 0.011 | 0.024 | 0.032 | 0.002 | 0.009 | 0.021 | 0.001 | 0.007 | 0.009 | 0.000 | 0.002 | 0.016 | 0.000 | 0.004 | 0.003 | 0.000 | 0.000 | 0.009 | 0.000 | 0.003 |
| Peque | 0.369 | 0.022 | 0.033 | 0.021 | 0.001 | 0.008 | 0.014 | 0.000 | 0.003 | 0.010 | 0.000 | 0.003 | 0.035 | 0.002 | 0.009 | 0.007 | 0.000 | 0.003 | 0.020 | 0.002 | 0.010 |
| Medellin | 0.664 | 0.018 | 0.030 | 0.077 | 0.006 | 0.018 | 0.023 | 0.001 | 0.005 | 0.011 | 0.000 | 0.005 | 0.018 | 0.001 | 0.005 | 0.004 | 0.000 | 0.001 | 0.006 | 0.000 | 0.001 |
| Cundinamarca | 0.472 | 0.021 | 0.033 | 0.007 | 0.000 | 0.003 | 0.029 | 0.001 | 0.009 | 0.013 | 0.000 | 0.004 | 0.043 | 0.004 | 0.015 | 0.015 | 0.001 | 0.006 | 0.012 | 0.000 | 0.002 |
| Pasto | 0.391 | 0.029 | 0.039 | 0.018 | 0.003 | 0.013 | 0.052 | 0.007 | 0.019 | 0.021 | 0.001 | 0.006 | 0.015 | 0.000 | 0.003 | 0.004 | 0.000 | 0.001 | 0.026 | 0.002 | 0.009 |
| Paposo | 0.391 | 0.010 | 0.022 | 0.003 | 0.000 | 0.001 | 0.014 | 0.001 | 0.005 | 0.010 | 0.000 | 0.002 | 0.013 | 0.000 | 0.004 | 0.004 | 0.000 | 0.001 | 0.022 | 0.002 | 0.010 |
| Quetalmahue | 0.416 | 0.006 | 0.017 | 0.001 | 0.000 | 0.000 | 0.003 | 0.000 | 0.000 | 0.006 | 0.000 | 0.002 | 0.004 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 |
| Salta | 0.232 | 0.044 | 0.048 | 0.004 | 0.000 | 0.001 | 0.016 | 0.000 | 0.004 | 0.018 | 0.002 | 0.010 | 0.036 | 0.006 | 0.017 | 0.007 | 0.000 | 0.002 | 0.033 | 0.003 | 0.013 |
| Tucuman | 0.649 | 0.034 | 0.042 | 0.024 | 0.001 | 0.006 | 0.017 | 0.000 | 0.005 | 0.027 | 0.003 | 0.013 | 0.015 | 0.000 | 0.004 | 0.010 | 0.000 | 0.004 | 0.013 | 0.001 | 0.006 |
| Catamarca | 0.537 | 0.007 | 0.022 | 0.005 | 0.000 | 0.001 | 0.050 | 0.004 | 0.018 | 0.013 | 0.000 | 0.004 | 0.030 | 0.002 | 0.011 | 0.016 | 0.000 | 0.006 | 0.012 | 0.000 | 0.003 |
| RGS | 0.703 | 0.024 | 0.034 | 0.083 | 0.014 | 0.027 | 0.011 | 0.000 | 0.002 | 0.006 | 0.000 | 0.001 | 0.014 | 0.000 | 0.003 | 0.003 | 0.000 | 0.000 | 0.011 | 0.000 | 0.003 |
| Population | Kaqchikel | | | Guaymi | | | Cabecar | | | Embera | | | Waunana | | | Zenu | | | Arhuaco | | |
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.033 | 0.006 | 0.018 | 0.009 | 0.000 | 0.003 | 0.006 | 0.000 | 0.002 | 0.005 | 0.000 | 0.001 | 0.008 | 0.000 | 0.002 | 0.012 | 0.001 | 0.008 | 0.005 | 0.000 | 0.002 |
| Oriente | 0.014 | 0.000 | 0.003 | 0.017 | 0.001 | 0.005 | 0.005 | 0.000 | 0.001 | 0.014 | 0.001 | 0.005 | 0.032 | 0.003 | 0.012 | 0.006 | 0.000 | 0.002 | 0.010 | 0.000 | 0.003 |
| CVCR | 0.011 | 0.000 | 0.003 | 0.008 | 0.000 | 0.004 | 0.006 | 0.000 | 0.001 | 0.047 | 0.003 | 0.012 | 0.012 | 0.000 | 0.003 | 0.020 | 0.001 | 0.006 | 0.014 | 0.000 | 0.004 |
| Peque | 0.008 | 0.000 | 0.001 | 0.014 | 0.000 | 0.004 | 0.009 | 0.000 | 0.002 | 0.095 | 0.008 | 0.020 | 0.023 | 0.001 | 0.008 | 0.051 | 0.004 | 0.014 | 0.044 | 0.003 | 0.013 |
| Medellin | 0.012 | 0.000 | 0.003 | 0.007 | 0.000 | 0.002 | 0.007 | 0.000 | 0.002 | 0.016 | 0.000 | 0.005 | 0.007 | 0.000 | 0.002 | 0.013 | 0.000 | 0.004 | 0.011 | 0.000 | 0.003 |
| Cundinamarca | 0.015 | 0.001 | 0.006 | 0.010 | 0.000 | 0.004 | 0.005 | 0.000 | 0.001 | 0.057 | 0.003 | 0.012 | 0.029 | 0.002 | 0.010 | 0.021 | 0.001 | 0.009 | 0.017 | 0.000 | 0.003 |
| Pasto | 0.017 | 0.000 | 0.003 | 0.006 | 0.000 | 0.002 | 0.005 | 0.000 | 0.001 | 0.011 | 0.001 | 0.005 | 0.013 | 0.000 | 0.004 | 0.010 | 0.000 | 0.002 | 0.009 | 0.000 | 0.003 |
| Paposo | 0.008 | 0.000 | 0.003 | 0.003 | 0.000 | 0.000 | 0.004 | 0.000 | 0.001 | 0.004 | 0.000 | 0.001 | 0.005 | 0.000 | 0.001 | 0.005 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 |
| Quetalmahue | 0.003 | 0.000 | 0.000 | 0.006 | 0.000 | 0.001 | 0.003 | 0.000 | 0.000 | 0.002 | 0.000 | 0.000 | 0.003 | 0.000 | 0.000 | 0.003 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 |
| Salta | 0.013 | 0.000 | 0.003 | 0.004 | 0.000 | 0.000 | 0.003 | 0.000 | 0.001 | 0.012 | 0.000 | 0.004 | 0.018 | 0.001 | 0.007 | 0.008 | 0.000 | 0.003 | 0.006 | 0.000 | 0.001 |
| Tucuman | 0.012 | 0.000 | 0.003 | 0.007 | 0.000 | 0.003 | 0.007 | 0.000 | 0.004 | 0.005 | 0.000 | 0.001 | 0.012 | 0.000 | 0.005 | 0.005 | 0.000 | 0.002 | 0.004 | 0.000 | 0.001 |
| Catamarca | 0.010 | 0.000 | 0.004 | 0.007 | 0.000 | 0.002 | 0.011 | 0.000 | 0.004 | 0.009 | 0.000 | 0.003 | 0.010 | 0.000 | 0.003 | 0.005 | 0.000 | 0.002 | 0.005 | 0.000 | 0.001 |
| RGS | 0.009 | 0.000 | 0.002 | 0.006 | 0.000 | 0.002 | 0.003 | 0.000 | 0.000 | 0.003 | 0.000 | 0.000 | 0.008 | 0.000 | 0.003 | 0.003 | 0.000 | 0.001 | 0.005 | 0.000 | 0.002 |

TableS5 (continued)

| Population | Kogi | | | Inga | | | Quechua | | | Aymara | | | Huilliche | | | Wayuu | | | Piapoco | | |
|--------------|-------|-------|-------|-------|-------|-------|---------|-------|-------|--------|-------|-------|-----------|-------|-------|-------|-------|-------|---------|-------|-------|
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.004 | 0.000 | 0.001 | 0.024 | 0.002 | 0.011 | 0.039 | 0.006 | 0.018 | 0.025 | 0.002 | 0.009 | 0.023 | 0.001 | 0.008 | 0.010 | 0.000 | 0.003 | 0.012 | 0.000 | 0.004 |
| Oriente | 0.011 | 0.000 | 0.003 | 0.025 | 0.003 | 0.012 | 0.036 | 0.003 | 0.013 | 0.024 | 0.001 | 0.006 | 0.027 | 0.003 | 0.012 | 0.011 | 0.000 | 0.004 | 0.022 | 0.002 | 0.011 |
| CVCR | 0.012 | 0.000 | 0.004 | 0.007 | 0.000 | 0.001 | 0.018 | 0.001 | 0.005 | 0.015 | 0.000 | 0.004 | 0.013 | 0.001 | 0.006 | 0.006 | 0.000 | 0.001 | 0.021 | 0.001 | 0.006 |
| Peque | 0.018 | 0.000 | 0.004 | 0.009 | 0.000 | 0.002 | 0.017 | 0.001 | 0.007 | 0.017 | 0.000 | 0.005 | 0.008 | 0.000 | 0.002 | 0.021 | 0.002 | 0.009 | 0.145 | 0.020 | 0.031 |
| Medellin | 0.010 | 0.000 | 0.004 | 0.018 | 0.001 | 0.007 | 0.012 | 0.000 | 0.002 | 0.009 | 0.000 | 0.002 | 0.007 | 0.000 | 0.002 | 0.006 | 0.000 | 0.001 | 0.026 | 0.001 | 0.006 |
| Cundinamarca | 0.019 | 0.001 | 0.006 | 0.022 | 0.001 | 0.005 | 0.018 | 0.000 | 0.004 | 0.035 | 0.003 | 0.012 | 0.028 | 0.001 | 0.008 | 0.017 | 0.000 | 0.004 | 0.045 | 0.003 | 0.014 |
| Pasto | 0.013 | 0.000 | 0.003 | 0.178 | 0.016 | 0.029 | 0.042 | 0.005 | 0.017 | 0.043 | 0.003 | 0.013 | 0.029 | 0.002 | 0.011 | 0.008 | 0.000 | 0.001 | 0.013 | 0.000 | 0.003 |
| Paposo | 0.006 | 0.000 | 0.002 | 0.021 | 0.001 | 0.008 | 0.030 | 0.003 | 0.012 | 0.309 | 0.066 | 0.057 | 0.082 | 0.007 | 0.019 | 0.006 | 0.000 | 0.001 | 0.011 | 0.000 | 0.003 |
| Quetalmahue | 0.002 | 0.000 | 0.000 | 0.003 | 0.000 | 0.001 | 0.007 | 0.000 | 0.002 | 0.004 | 0.000 | 0.001 | 0.498 | 0.005 | 0.017 | 0.003 | 0.000 | 0.001 | 0.003 | 0.000 | 0.000 |
| Salta | 0.006 | 0.000 | 0.001 | 0.022 | 0.001 | 0.005 | 0.152 | 0.020 | 0.032 | 0.313 | 0.056 | 0.054 | 0.021 | 0.000 | 0.004 | 0.012 | 0.000 | 0.004 | 0.016 | 0.000 | 0.004 |
| Tucuman | 0.008 | 0.000 | 0.003 | 0.009 | 0.000 | 0.005 | 0.047 | 0.004 | 0.015 | 0.023 | 0.001 | 0.008 | 0.030 | 0.003 | 0.013 | 0.009 | 0.000 | 0.002 | 0.014 | 0.000 | 0.004 |
| Catamarca | 0.004 | 0.000 | 0.001 | 0.010 | 0.000 | 0.004 | 0.124 | 0.016 | 0.033 | 0.041 | 0.004 | 0.018 | 0.018 | 0.001 | 0.007 | 0.014 | 0.000 | 0.005 | 0.012 | 0.000 | 0.003 |
| RGS | 0.011 | 0.001 | 0.006 | 0.006 | 0.000 | 0.001 | 0.025 | 0.002 | 0.011 | 0.011 | 0.000 | 0.003 | 0.008 | 0.000 | 0.002 | 0.015 | 0.001 | 0.006 | 0.015 | 0.000 | 0.004 |

| Population | Ticuna (Arara) | | | Ticuna (Tarapaca) | | | Karitiana | | | Surui | | | Guarani | | | Ache | | | Kaingang | | |
|--------------|----------------|-------|-------|-------------------|-------|-------|-----------|-------|-------|-------|-------|-------|---------|-------|-------|-------|-------|-------|----------|-------|-------|
| | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. | Mean | Var. | s.e. |
| Mexico City | 0.005 | 0.000 | 0.001 | 0.010 | 0.001 | 0.006 | 0.005 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.003 | 0.000 | 0.000 | 0.010 | 0.000 | 0.003 | 0.009 | 0.000 | 0.004 |
| Oriente | 0.005 | 0.000 | 0.001 | 0.011 | 0.001 | 0.006 | 0.008 | 0.000 | 0.003 | 0.004 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.009 | 0.000 | 0.002 | 0.008 | 0.000 | 0.004 |
| CVCR | 0.004 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.003 | 0.000 | 0.000 | 0.004 | 0.000 | 0.001 | 0.002 | 0.000 | 0.001 | 0.006 | 0.000 | 0.001 | 0.004 | 0.000 | 0.002 |
| Peque | 0.009 | 0.000 | 0.002 | 0.011 | 0.000 | 0.004 | 0.005 | 0.000 | 0.001 | 0.005 | 0.000 | 0.001 | 0.002 | 0.000 | 0.000 | 0.009 | 0.000 | 0.003 | 0.005 | 0.000 | 0.001 |
| Medellin | 0.011 | 0.000 | 0.005 | 0.005 | 0.000 | 0.001 | 0.006 | 0.000 | 0.001 | 0.002 | 0.000 | 0.000 | 0.003 | 0.000 | 0.001 | 0.008 | 0.000 | 0.004 | 0.002 | 0.000 | 0.000 |
| Cundinamarca | 0.014 | 0.001 | 0.007 | 0.006 | 0.000 | 0.001 | 0.006 | 0.000 | 0.001 | 0.008 | 0.000 | 0.002 | 0.003 | 0.000 | 0.001 | 0.025 | 0.001 | 0.009 | 0.009 | 0.000 | 0.004 |
| Pasto | 0.014 | 0.000 | 0.004 | 0.007 | 0.000 | 0.002 | 0.023 | 0.001 | 0.006 | 0.007 | 0.000 | 0.003 | 0.003 | 0.000 | 0.001 | 0.020 | 0.001 | 0.008 | 0.003 | 0.000 | 0.001 |
| Paposo | 0.004 | 0.000 | 0.001 | 0.004 | 0.000 | 0.001 | 0.009 | 0.000 | 0.004 | 0.004 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.017 | 0.002 | 0.009 | 0.003 | 0.000 | 0.001 |
| Quetalmahue | 0.003 | 0.000 | 0.001 | 0.002 | 0.000 | 0.000 | 0.002 | 0.000 | 0.000 | 0.004 | 0.000 | 0.001 | 0.003 | 0.000 | 0.001 | 0.005 | 0.000 | 0.002 | 0.003 | 0.000 | 0.001 |
| Salta | 0.012 | 0.000 | 0.003 | 0.004 | 0.000 | 0.001 | 0.008 | 0.000 | 0.003 | 0.005 | 0.000 | 0.001 | 0.006 | 0.000 | 0.002 | 0.009 | 0.000 | 0.002 | 0.004 | 0.000 | 0.001 |
| Tucuman | 0.010 | 0.000 | 0.003 | 0.005 | 0.000 | 0.003 | 0.003 | 0.000 | 0.001 | 0.006 | 0.000 | 0.002 | 0.004 | 0.000 | 0.001 | 0.023 | 0.003 | 0.012 | 0.003 | 0.000 | 0.001 |
| Catamarca | 0.007 | 0.000 | 0.002 | 0.005 | 0.000 | 0.002 | 0.005 | 0.000 | 0.001 | 0.007 | 0.000 | 0.003 | 0.003 | 0.000 | 0.001 | 0.022 | 0.001 | 0.009 | 0.005 | 0.000 | 0.002 |
| RGS | 0.005 | 0.000 | 0.001 | 0.005 | 0.000 | 0.001 | 0.005 | 0.000 | 0.001 | 0.006 | 0.000 | 0.002 | 0.003 | 0.000 | 0.001 | 0.013 | 0.000 | 0.005 | 0.006 | 0.000 | 0.002 |

Supplementary Figure S1

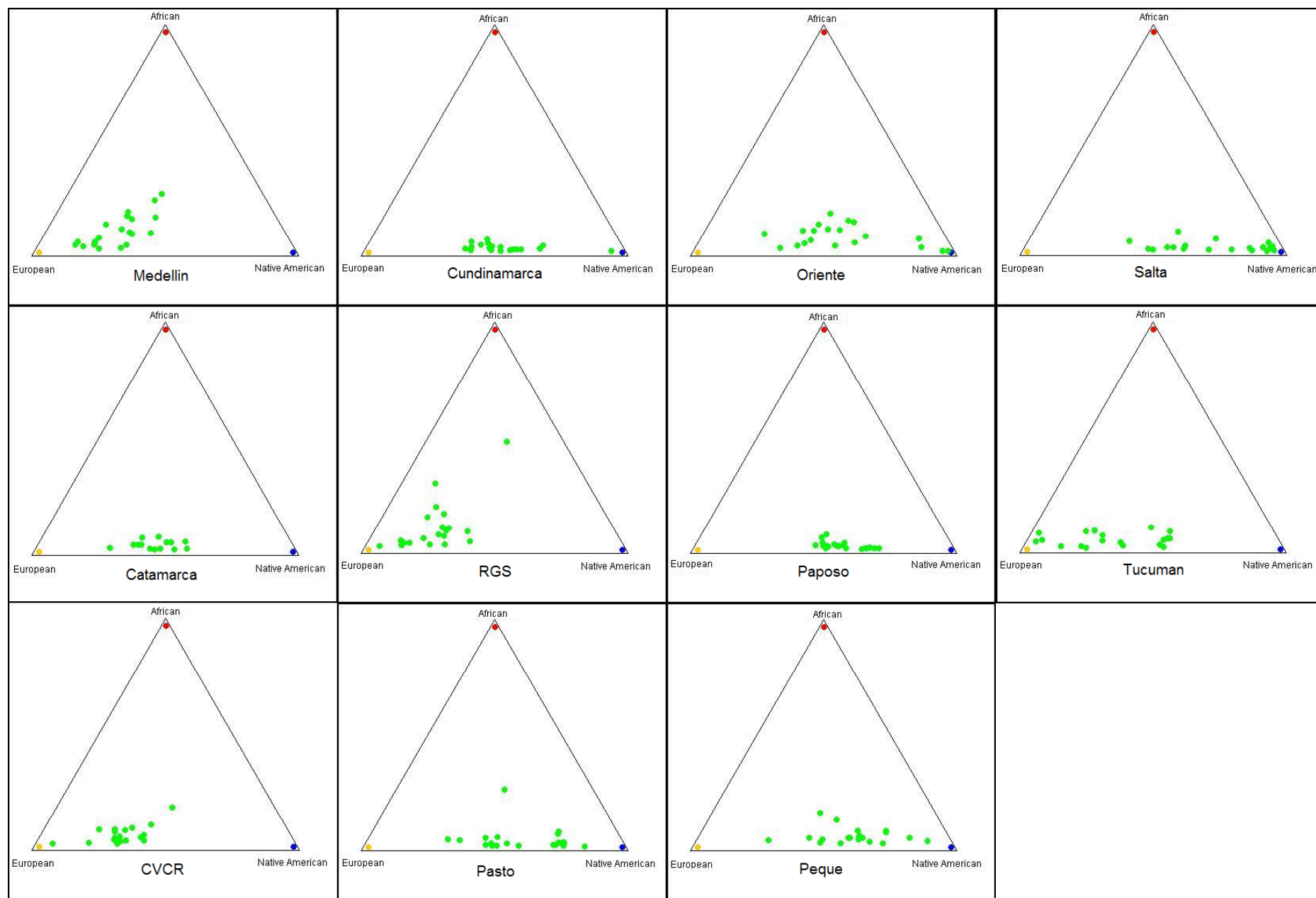


FIGURE LEGEND

FigureS1 Distribution of admixture estimates for individuals from 11 Mestizo populations. The position of each green dot on the triangle plot indicates the proportion of European, Native American and African ancestry estimated for each individual in the population.