AMERICAN JOURNAL OF HUMAN BIOLOGY 23:299–304 (2011)

Original Research Article

X-Chromosomal Genetic Diversity and Linkage Disequilibrium Patterns in Amerindians and Non-Amerindian Populations

CARLOS EDUARDO G. AMORIM,[1] SIJIA WANG,[2] ANDREA R. MARRERO,[1] FRANCISCO M. SALZANO,[1] ANDRÉS RUIZ-LINARES,[2] AND MARIA CÁTIRA BORTOLINI[1]*
[1]*Programa de Pós-Graduação em Genética e Biologia Molecular and Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970, Porto Alegre, RS, Brazil*
[2]*The Galton Laboratory, Department of Biology, University College London, London, United Kingdom*

Objectives: We report X-chromosomal linkage disequilibrium (LD) patterns in Amerindian (Kogi, Wayuu, and Zenu) and admixed Latin American (Central Valley of Costa Rica and Southern Brazilian Gaucho) populations.

Methods: Short tandem repeats (STRs) widespread along the X-chromosome were investigated in 132 and 124 chromosomes sampled from the Amerindian tribes and the admixed Latin American populations, respectively. Diversity indexes (gene diversity and average numbers of alleles per locus) were estimated for each population and the level of LD was inferred with an exact test.

Results: The Amerindian populations presented lower genetic diversity and a higher proportion of loci in LD than the admixed ones. Two haplotype blocks were identified in the X-chromosome, both restricted to the Amerindians. The first involved DXS8051 and DXS7108 in Xp22.22 and Xp22.3, while the second found only among the Kogi, included eight loci in a region between Xp11.4 and Xq21.1.

Conclusions: In accordance to previous work done with other populations, human isolates, such as Amerindian tribes, seem to be an optimal choice for the implementation of association studies due to the wide extent of LD which can be found in their gene pool. On the other hand, the low proportion of loci in LD found in both admixed populations studied here could be explained by events related to their history and similarities between the allele frequencies in the parental stocks. Am. J. Hum. Biol. 23:299–304, 2011.

Understanding patterns of linkage disequilibrium (LD), i.e., the nonrandom association of alleles at two or more loci, is the basis for human gene mapping and for the design of association studies (reviewed by Slatkin, 2008). It also provides information on many aspects of population history and evolution, such as the occurrence of natural selection, past demographic events, gene flow, population structure, and breeding system (Ardlie et al., 2002; Pfaff et al., 2001; Reich et al., 2001; Slatkin, 2008). LD serves as the theoretical foundation for association mapping—a marker and a functional locus need to be in LD for the association to be detected. Consequently, populations in high LD are the best choices for designing gene mapping strategies, as the number of markers employed for the identification of an associated allele can be reduced (Ardlie et al., 2002; Slatkin, 2008). In particular, small isolates present the desired profile for gene mapping studies because they commonly have high levels of LD caused by drift effects (Katoh et al., 2002; Varilo and Peltonen, 2004; Service et al., 2006). Admixed populations are also of interest, because their gene pool was formed by relatively recent events involving distinct parental stocks, which may result in long-range LD (Pfaff et al., 2001).

Many Amerindian and Latin American populations present these two characteristics and therefore seem to be optimal choices for gene mapping and for the implementation of association studies. Notwithstanding linkage disequilibrium investigations among them are scarce. We therefore developed an investigation design, which is an extension of previous and recent investigations on X-STR LD in distinct pools of Amerindian and non-Amerindian samples (Leite et al., 2009; Wang et al., 2010). In our study three relatively isolated Native American populations and two non-native communities were tested for a fast-evolving genetic system, with the following questions in mind: (a) since genetic variability clearly influences the power to detect LD, do the populations investigated here show differences in this characteristic that could influence the patterns obtained? (b) In the populations tested what factors could be more adequate to explain the LD levels? and (c) Could haplotype blocks be distinguished that would provide new tools for the investigation of the phylogeography of these populations? The results indicated that the Amerindian tribes presented less genetic diversity and wider LD extent than the remaining populations, where LD was virtually absent. This pattern is most likely associated to the evolutionary history of these populations and also to the genetic systems chosen for the analyses.

SUBJECTS AND METHODS

Population description and data source

New data for 47 X-chromosomal short tandem repeats (STRs) were generated for three South Amerindian populations from Colombia. Seventeen of these markers were

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

Contract grant sponsors: Institutos do Milênio and Apoio a Núcleos de Excelência Programs; Conselho Nacional de Desenvolvimento Científico e Tecnológico; Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul.

Carlos Eduardo G. Amorim and Sijia Wang contributed equally to this work.

*Correspondence to: Maria Cátira Bortolini, Programa de Pós-Graduação em Genética e Biologia Molecular and Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970, Porto Alegre, RS, Brazil.
E-mail: maria.bortolini@ufrgs.br

Received 5 April 2010; Revision received 8 September 2010; Accepted 14 September 2010

DOI 10.1002/ajhb.21110

Published online 3 February 2011 in Wiley Online Library (wileyonlinelibrary.com).

TABLE 1. Geographical location of the five populations analyzed

Population	Ethnic composition	Country	Geographical coordinates	
Central Valley of Costa Rica	Admixed	Costa Rica	9°56′N	84°05′W
Southern Brazilian Gaucho	Admixed	Brazil	31°00′S	54°00′W
Kogi	Amerindian	Colombia	11°00′N	74°00′W
Wayuu	Amerindian	Colombia	9°00′N	75°00′W
Zenu	Amerindian	Colombia	11°00′N	73°00′W

also investigated in a sample of inhabitants of the Pampa region, which corresponds to parts of Argentina, Uruguay, and southern Brazil. These people are known as "Gaucho" (Marrero et al., 2007).

Our Amerindian sample comprises 132 chromosomes sampled from Kogi (Chibchan-Paezan linguistic stock), Wayuu (Equatorial-Tucanoan linguistic stock), and Zenu (currently speaking Spanish; Mesa et al.,2000). For each population, 13 women and 18 men have been genotyped. The Gaucho genotyped here ($N = 70$) are a subset of the male sample collected in the Brazilian Pampa region in the cities of Bagé and Alegrete in the Brazilian state of Rio Grande do Sul previously described by Marrero et al. (2007).

A fifth population, located in the Central Valley of Costa Rica (CVCR), was included in the analysis. This admixed population was previously described by Service et al. (2001; data kindly provided) and, in our work we analyzed the 54 male individuals. Briefly, CVCR was founded by Spaniards and Amerindians in the 16th to 18th centuries, and the current population number is two to three million; they are relatively isolated from communities of the Pacific and Atlantic coastal regions.

Additional information about the populations studied is as follows: (a) the Kogi live in Sierra Nevada de Santa Marta, and have been relatively isolated from non-Amerindians (Zarante et al., 2000). A summary of their environment and culture, basically derived from the work of Gerardo Reichel-Dolmatoff performed in the 1940s, can be found in Wilson (1999); (b) Wayuu and Zenu represent the other Amerindian populations which, differently from Kogi, show signs of admixture with non-native groups during America's colonization (Mesa et al., 2000; Wang et al., 2007); and (c) the Gaucho population was basically formed at 18th century through intermarriages between the European colonizers (mainly Portuguese and Spaniard males) and Amerindian females. The African component was introduced later with the first slaves in the region (Marrero et al., 2007; Wang et al., 2008). Geographical location of each population can be found in Table 1.

Ethical approval for the present study was provided by the Brazilian National Ethics Commission (CONEP; resolution number: 1333/2002) and by ethics committees in the countries where the non-Brazilian samples were collected.

Genotyping procedure

Forty-seven STRs were genotyped for the Amerindian populations. This set of loci comprises DXS1039, DXS1216, plus the markers included in ABI Prism linkage Panels 28, 83 (except for DXS8088), 84, 85, and 86. The same information was compiled for the 54 CVCR male individuals studied earlier (Service et al., 2001). A subset of these loci, encompassing 17 markers of Panel 28 (except DXS8043), was typed for the Gaucho. Genotyping was performed according to the user's manual provided by the manufacturer (ABI Prism) using an ABI 3730xl sequencer and with GENEMAPPER v3.5. All the newly obtained genotypes (classified according to allele length) are included in the Supporting Information Dataset S1.

Statistical analysis

Phase v.2.1 (Stephens and Scheet, 2005; Stephens et al., 2001) was employed to resolve the haplotype phase of the 39 female Amerindians using default settings. The stepwise mutation model was chosen for the analysis. Males from these populations were randomly paired to form pseudo-diploid individuals and were used as known-phase individuals, as suggested by Phase's authors (Mathew Stephens, personal communication). To control for inference errors, three runs were performed for each sample and the three outputs were then compared. No differences were observed.

Allele frequencies and mean number of alleles per locus were estimated by direct count. The Arlequin package v.3.11 (Excoffier et al., 2005) was employed to calculate the average gene diversity ($\hat{H}$; i.e., the probability that two randomly chosen haplotypes are different in a sample) and to perform the LD analysis for all populations. For the LD estimates, Arlequin employs an exact test to check for non-random association of alleles at different loci. The output *P*-values for the association tests were then corrected following the Benjamini and Hochberg false discovery rate procedure (Benjamini and Hochberg, 1995). $\hat{H}$ and mean number of alleles per locus were compared across populations with Dunn's test ($\alpha = 0.05$) using the BioEstat 5.0 software (available at www.mamiraua.com.br) after significant differences in these statistics had been found by a Kruskal-Wallis test ($\alpha = 0.05$).

We then looked for haplotype blocks by examining sequences of two or more markers in a row with significant linkage disequilibrium. Further recombination analyses were conducted for each population with Phase v.2.1 (Li and Stephens, 2003; Crawford et al., 2004) with the presumed haplotype blocks to estimate their background recombination rate (ρ) in "per base pair" units. This software gives a ρ distribution of values generated by "*n*" successive runs, which we set to 100. Haplotype networks for both blocks were generated with the Network 4.5.1.0 package (Bandelt et al., 1999) using the Median Joining method, while the mean number of pairwise differences between haplotypes were estimated with the Arlequin package v.3.11 (Excoffier et al., 2005).

RESULTS

For the 47 X-chromosomal STRs, average number of alleles per locus and $\hat{H}$ were both significantly lower for the three Amerindian tribes as compared to CVCR (Table 2). When the Gaucho sample was considered, decreasing the number of markers to 17, a similar pattern was observed: the mean number of alleles per locus was lower for the Amerindians as compared to CVCR and Gaucho, and $\hat{H}$ was significantly different between Kogi and Gaucho (Table 3).

In the 17-loci dataset the Gaucho presented the lowest degree of LD, followed by CVCR and Zenu. The latter showed the lowest values in the 47-loci results. Kogi and Wayuu presented the highest proportion of loci in LD, but

TABLE 2. Genetic information for Amerindians and the Central Valley of Costa Rica (CVCR) based on the allelic distribution of 47 STRs[a]

Characteristic	Kogi	Wayuu	Zenu	CVCR
Average ± SD gene diversity ($\hat{H}$)	$0.51^{a} \pm 0.25$	$0.57^{a} \pm 0.28$	$0.59^{a} \pm 0.29$	$0.67^{b} \pm 0.33$
Average no. of alleles per locus	3.70^{a}	4.64^{ab}	5.15^{b}	6.87^{c}
Proportion of loci in LD before correction (%)	31.69	22.11	15.82	6.75
Proportion of loci in LD after correction (%)	15.65	3.89	2.59	1.76

[a]Values followed by the same letter present no statistically significant differences according to Dunn's test ($\alpha = 0.05$).

TABLE 3. Genetic information for Amerindian (Kogi, Wayuu, and Zenu) and non-Amerindian (CVCR and Gaucho) populations based on the allelic distribution of 17 STRs[a]

	Kogi	Wayuu	Zenu	CVCR	Gaucho
Average ± SD gene diversity ($\hat{H}$)	$0.52^{a} \pm 0.27$	$0.61^{ab} \pm 0.32$	$0.59^{abc} \pm 0.31$	$0.72^{bc} \pm 0.36$	$0.72^{c} \pm 0.37$
Average no. of alleles per locus	3.23^{a}	4.47^{a}	4.82^{a}	7.06^{b}	8.18^{b}
Proportion of loci in LD before correction (%)	28.33	35.29	23.53	6.62	4.41
Proportion of loci in LD after correction (%)	14.17	22.79	8.09	0	0

[a]Values followed by the same letter present no statistically significant differences according to Dunn's test ($\alpha = 0.05$).

TABLE 4. DXS8051 and DXS7108 haplotype distribution in three Colombian Amerindian populations[a]

			Frequencies			
Haplotype	DXS8051	DXS7108	Kogi	Wayuu	Zenu	Total
h1	112	252	–	–	1	1
h2	114	260	1	–	1	2
h3	116	252	–	3	1	4
h4	116	262	–	2	–	2
h5	118	252	5	1	3	9
h6	118	254	3	–	–	3
h7	118	256	–	–	3	3
h8	118	260	5	2	4	11
h9	118	262	4	1	9	14
h10	120	252	–	–	2	2
h11	120	260	–	–	1	1
h12	122	252	1	–	–	1
h13	122	260	8	6	–	14
h14	122	262	–	7	–	7
h15	124	252	5	5	2	12
h16	124	260	–	6	–	6
h17	124	262	2	–	–	2
h18	124	262	–	2	–	2
h19	126	252	4	–	–	4
h20	126	256	1	–	–	1
h21	126	260	–	–	7	7
h22	126	262	1	–	–	1
h23	128	252	–	4	–	4
h24	128	258	–	–	1	1
h25	128	260	2	–	2	4
h26	128	262	1	–	–	1
h27	128	264	–	–	2	2
h28	130	260	1	–	–	1
h29	132	252	–	–	2	2
h30	132	262	–	5	–	5
h31	134	252	–	–	2	2
h32	134	262	–	–	1	1
Alelles	12	7				132

[a]Alleles are classified according to allele length.

showed distinct ranks when the two different sets of markers were employed (Tables 2 and 3).

Two haplotype blocks of interest were identified, both in Amerindians. The first is common to all three Amerindian populations. It is defined by two loci (DXS8051 and DXS7108) and is located from Xp22.22 to Xp22.3, involving a region of 1,080 kb, with a genetic distance estimated as 2.2 cM (http://www.appliedbiosystems.com). Average $\hat{H}$ for DXS8051 and DXS7108 in Amerindians was estimated as 0.81 and 0.71 respectively, while the recombination rates between these loci were 0.0005, 0.0013, and 0.0005 for the Kogi, Wayuu, and Zenu populations, respectively. The haplotype network presented several reticulations and therefore is not shown. Possible causes for this level of reticulation include back mutation, homoplasy, and statistical limitations of the methodological procedures used to generate the networks. As given in Table 4, the most common haplotypes in the whole Amerindian sample were h9 (118-262) and h13 (122-260); one of them was also the most common haplotype in the Zenu and Kogi, but the haplotype with the highest frequency among the Wayuu was h14 (122-262). Kogi, Wayuu, and Zenu presented eight, six, and ten private haplotypes respectively.

The Kogi presented many haplotype blocks defined by few markers. However, one of these blocks deserves mention due to its size. It includes eight loci (DXS993, DXS8080, DXS8083, DXS1055, DXS1039, DXS991, DXS1216, and DXS986) with a background recombination rate (ρ) estimated at 6.3×10^{-5}. The first and last markers in this block are 11 kb apart from each other, and delimit a region between Xp11.4 and Xq21.1, therefore including the centromere. All these loci except DXS1039, presented higher $\hat{H}$ than the overall mean. DXS1039 presents the lowest proportion of significant LD *P*-values and very low $\hat{H}$ value. The network with this eight-loci haplotype block is shown in Figure 1, while the haplotype frequencies are given in Table 5. Haplotype h8 (275-82-183-152-189-331-250-175) was the most frequent (22%), followed by h7 (275-82-183-150-189-333-250-175; 14%). Mean number of pairwise differences for this haplotype block was estimated as 4.73 ± 2.36 and $\hat{H}$ as 0.59 ± 0.22.

DISCUSSION

Many aspects of population evolutionary and demographic trajectories can affect LD patterns. It is known, for instance, that natural selection, drift, and gene flow can raise the number of associated alleles in a sample, but the exact way these mechanisms affect linkage is not fully understood. In our study, we examined five distinct populations with different evolutionary or demographic histories to better understand the forces that can lead to distinct patterns of LD in such groups. The Amerindian populations presented lower intra-population genetic diversity

and higher extent of LD than the other two (CVCR and Gaucho) groups. While the pattern observed in the three Amerindian populations can be explained according to the evolutionary and demographic dynamics of such groups in the Americas, the low level of LD observed in CVCR and Gaucho can be associated with their admixture dynamics and the set of markers employed.

The first populations that arrived in the Americas were initially subjected to a moderate population bottleneck during the entry into the continent (Fagundes et al., 2008). Additionally, they experienced successive posterior population bottlenecks due to the dramatic effects of the fission–fusion events of village propagation and tribalization (Bortolini et al., 2003; Neel and Salzano, 1967). As a consequence of these demographic and evolutionary phenomena, genetic drift became important. Presently, Amerindian populations show distinct genetic characteristics when compared to those of other continents, such as the highest interpopulation divergence and the lowest intrapopulation diversity when a large number of autosomal fast-evolving markers are considered (Wang et al., 2007). Additionally, private Native American alleles have emerged (Schroeder et al., 2007, 2009; Acuña-Alonzo et al., 2010).

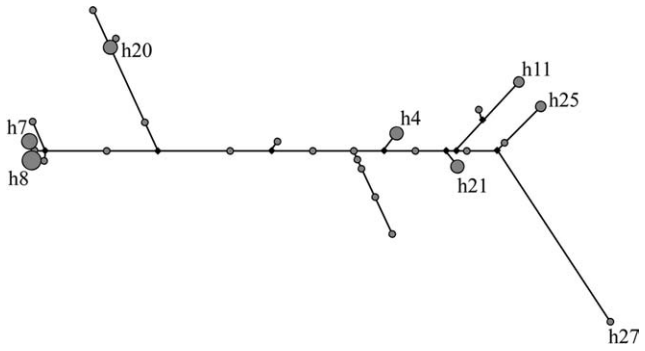

Fig. 1. Kogi's haplotype network considering eight X-chromosome loci. Grey circles represent the haplotypes and their sizes are proportional to haplotype frequencies. Black diamonds represent median vectors; h28 was not included in this analysis because its inclusion would lead to reticulation.

It is known that population bottlenecks influence LD, since genetic drift may easily cause haplotypes loss, generally resulting in increased LD (Slatkin, 2008). Moreover, small effective sizes, which are a characteristic of many Native American populations, may decrease the occurrence of recombination and consequently may also raise LD levels. Thus, the lower genetic diversity and the higher level of LD observed in the Native populations, as compared to the CVCR and Gaucho samples, might be associated with these phenomena. Similar results were obtained in previous independent studies involving South Amerindian (Leite et al., 2009; Wang et al., 2010) and other small and isolated populations such as the Khoton from Mongolia (Katoh et al., 2002), the Saami (Laan and Pääbo, 1997), and Kuusamo (Varilo et al., 2000) from Finland.

The lower genetic diversity of the Kogi in comparison to Wayuu and Zenu might be associated with the fact that they present a higher degree of isolation, with less influence of admixture with nonautochthonous groups (Mesa et al., 2000; Zarante et al., 2000; Wang et al., 2007). This fact, in addition to their reduced population size, may also be responsible for the Kogi's increased LD and for the occurrence of the private long-range haplotype block spanning Xp11.4 and Xq21.1.

TABLE 5. Haplotype distribution considering eight loci in the Kogi, a Colombian Amerindian population[a]

Haplotype	DXS993	DXS8080	DXS8083	DXS1055	DXS1039	DXS991	DXS1216	DXS986	Frequency
h1	273	96	181	150	189	325	254	175	1
h2	273	96	181	150	189	327	254	167	1
h3	273	96	181	150	189	327	254	175	1
h4	273	96	185	154	189	327	254	177	3
h5	273	98	181	150	189	325	254	175	1
h6	275	82	181	156	189	335	250	175	1
h7	275	82	183	150	189	333	250	175	4
h8	275	82	183	152	189	331	250	175	6
h9	275	82	183	152	189	333	250	175	1
h10	275	82	183	154	189	333	254	167	1
h11	275	96	181	150	189	333	260	175	2
h12	275	96	181	154	189	331	254	177	1
h13	275	96	181	154	189	331	260	167	1
h14	277	82	183	154	189	333	250	175	1
h15	277	96	181	150	189	327	260	167	1
h16	277	100	181	150	189	325	254	167	1
h17	277	100	181	150	189	325	254	175	1
h18	281	82	181	154	189	331	260	167	1
h19	281	82	185	152	189	327	260	175	1
h20	281	82	185	154	189	327	260	175	3
h21	281	96	181	154	189	327	254	177	3
h22	281	96	181	154	189	333	254	167	1
h23	281	96	181	154	189	333	254	175	1
h24	281	96	183	152	189	331	260	175	1
h25	281	96	185	154	189	327	254	167	2
h26	289	82	185	154	189	327	260	175	1
h27	295	90	175	148	187	331	250	169	1
h28	275	96	181	154	189	331	254	175	
Alleles	6	5	4	5	2	5	3	4	43

[a]Alleles are classified according to allele length.

On the other hand, the low proportion of loci in LD in the CVCR and Gaucho samples deserves additional attention. These populations present a complex pattern of admixture for the X-chromosome, involving mainly Amerindians and Europeans (the African contribution is also detected, but at lower proportions). The relative ancestral contribution to their X-chromosomes is not very different [in percentages, CVCR, European (E): 40; Amerindian (Am): 42; African (Af): 18; Gaucho, E: 47; Am: 31; Af: 22; (Wang et al., 2008)].

Different admixture models can create distinct LD patterns. Long (1991) proposed two extreme models of admixture: the hybrid-isolation (HI) and the continuous-gene-flow (CGF) models. In the HI model, admixture occurs in a single generation producing long-range LD, which progressively decays in each successive generation as a result of independent assortment and recombination between loci. In the CGF model, admixture occurs at a steady rate in every generation, and the amount of LD increases during the first few generations as continual admixture generates more disequilibrium than is broken down by independent assortment and recombination. After a few generations, the amount of LD begins to decay, although at a much slower rate than that observed in the HI model (Pfaff et al., 2001). The last authors, using a simulation approach, observed that under the HI model and considering unlinked loci (genetic distances between loci greater than 5 cM), only five generations are necessary after admixture for LD to decay to values close to zero (Pfaff et al., 2001). Applying this theory to our work, based on age of admixture estimates generated elsewhere (6.57 generations for the Gaucho and 14.42 for CVCR; Wang et al., 2008), the observed low proportion of loci in LD (Tables 2 and 3) could be explained assuming the hybrid-isolation model for unlinked markers for both admixed populations studied here. Furthermore, similarities between allele frequencies in the parental stocks could also be related to the pattern observed to Gaucho and CVCR populations.

The comparison between X-chromosomal diversity and LD between Amerindian and admixed groups was already examined in a previous investigation (Leite et al., 2009). This work shows some similarities to ours, with lower locus diversity, lower number of alleles per site, and higher LD among Amerindians as compared to the admixed populations. The diversity indices, however, are slightly lower in our work, indicating that the set of markers used by Leite et al. (2009) is more variable than ours, which may also be linked to the wider extent of LD detected in our study. Mention should also be made of the differences between the correction procedures and methods for LD detection used in both studies, which may have influenced the LD comparison.

Finally, the two haplotype blocks observed in the Amerindian X-chromosomes could be of potential interest, since LD analysis can also bring new insights into human evolutionary studies. It is reasonable to suggest that the most frequent haplotype observed in a specific population is the founder haplotype. However, the nature of the Amerindian evolutionary history, as mentioned before, strongly influenced by genetic drift due to isolation and fragmentation, prevents the identification of the Native American founder haplotype considering the *DXS8051* and *DXS7108* loci (Table 4). On the other hand, the network of Figure 1, based on eight loci (DXS993, DXS8080, DXS8083, DXS1055, DXS1039, DXS991, DXS1216, and DXS986), shows no evident signal of diversification from a founder haplotype, and the variable haplotype frequencies can also be associated with a scenario where genetic drift has a strong influence.

Despite these uncertainties, the two blocks reflect low-recombination regions, and it is possible to suggest that both, or at least parts of them, could be found in other Native American populations. In this context it is worth mentioning that a haplotype block between Xq13.3 and Xq21.3 was found in an investigation of 51 different populations, including South Amerindians (Santos-Lopes et al., 2007). In our analysis, only one marker (DXS1196) was included in this region, and therefore no appropriate comparison can be made between their findings and our data.

CONCLUSIONS

Now it is possible to answer the questions raised in the introduction. The Amerindian populations clearly show less intrapopulation variability than the two others and only among them LD blocks were detected. In this case, random drift should be considered as the most important factor for this detection. On the other hand, events as described in the hybrid-isolation model for unlinked markers and similarities between the allele frequencies in the parental stocks could be responsible for the absence of LD in both admixed populations. The two observed haplotype blocks can provide valuable new tools for Amerindian X-chromosome phylogeopraphic investigations, and this possibility will be explored by us in future studies.

ACKNOWLEDGMENTS

The authors thank all the blood donors for their participation, as well as Rogério Schmidt Campos for the designing support, and Sidia Maria Callegari-Jacques, Carla Martins Lopes, and Nelson Jurandi Rosa Fagundes for statistical and methodological help.

LITERATURE CITED

Acuña-Alonzo V, Flores-Dorantes T, Kruit JK, Villarreal-Molina T, Arellano-Campos O, Hünemeier T, Moreno-Estrada A, Ortiz-López MG, Villamil-Ramírez H, León-Mimila P, Villalobos-Comparan M, Jacobo-Albavera L, Ramírez-Jiménez S, Sikora M, Zhang LH, Pape TD, Granados-Silvestre Mde A, Montufar-Robles I, Tito-Alvarez AM, Zurita-Salinas C, Bustos-Arriaga J, Cedillo-Barrón L, Gómez-Trejo C, Barquera-Lozano R, Vieira-Filho JP, Granados J, Romero-Hidalgo S, Huertas-Vázquez A, González-Martín A, Gorostiza A, Bonatto SL, Rodríguez-Cruz M, Wang L, Tusié-Luna T, Aguilar-Salinas CA, Lisker R, Moises RS, Menjivar M, Salzano FM, Knowler WC, Bortolini MC, Hayden MR, Baier LJ, Canizales-Quinteros S. 2010. A functional ABCA1 gene variant is associated with low HDL-cholesterol levels and shows evidence of positive selection in Native Americans. Hum Mol Genet 19:2877–2885.

Ardlie KG, Kruglyak L, Seielstad M. 2002. Patterns of linkage disequilibrium in the human genome. Nat Rev Genet 3:299–309.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16:37–48.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300.

Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SP, Bau CH, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT, Hill K, Hurtado AM, Castro-de-Guerra D, Torres MM, Groot H, Michalski R, Nymadawa P, Bedoya G, Bradman N, Labuda D, Ruiz-Linares A. 2003. Y-chromosome evidence for differing ancient demographic histories in the Americas. Am J Hum Genet 73:524–539.

Crawford DC, Bhangale T, Li N, Hellenthal G, Rieder MJ, Nickerson DA, Stephens M. 2004. Evidence for substantial fine-scale variation in recombination rates across the human genome. Nat Genet 36:700–706.

Excoffier L, Laval G, Schneider S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50.

Fagundes NJ, Kanitz R, Eckert R, Valls AC, Bogo MR, Salzano FM, Smith DG, Silva WA Jr, Zago MA, Ribeiro-dos-Santos AK, Santos SE, Petzl-Erler ML, Bonatto SL. 2008. Mitochondrial population genomics supports a single pre-Clovis origin with a coastal route for the peopling of the Americas. Am J Hum Genet 82:583–592.

Katoh T, Mano S, Ikuta T, Munkhbat B, Tounai K, Ando H, Munkhtuvshin N, Imanishi T, Inoko H, Tamiya G. 2002. Genetic isolates in East Asia: a study of linkage disequilibrium in the X chromosome. Am J Hum Genet 71:395–400.

Laan M, Pääbo S. 1997. Demographic history and linkage disequilibrium in human populations. Nat Genet 17:435–438.

Leite FP, Santos SE, Rodrigues EM, Callegari-Jacques SM, Demarchi DA, Tsuneto LT, Petzl-Erler ML, Salzano FM, Hutz MH. 2009. Linkage disequilibrium patterns and genetic structure of Amerindian and non-Amerindian Brazilian populations revealed by long-range X-STR markers. Am J Phys Anthropol 139:404–412.

Li N, Stephens M. 2003. Modelling linkage disequilibrium, and identifying recombination hotspots using SNP data. Genetics 165:2213–2233.

Long JC. 1991. The genetic structure of admixed populations. Genetics 127:417–428.

Marrero AR, Bravi C, Stuart S, Long JC, Pereira-das-Neves-Leite F, Kommers T, Carvalho CM, Pena SD, Ruiz-Linares A, Salzano FM, Bortolini MC. 2007. Pre- and post-Columbian gene and cultural continuity: the case of the Gaucho from southern Brazil. Hum Hered 64:160–171.

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

Neel JV, Salzano FM. 1967. Further studies on the Xavante Indians. X. Some hypotheses-generalizations resulting from these studies. Am J Hum Genet 19:554–574.

Pfaff CL, Parra EJ, Bonilla C, Hiester K, McKeigue PM, Kamboh MI, Hutchinson RG, Ferrell RE, Boerwinkle E, Shriver MD. 2001. Population structure in admixed populations: effect of admixture dynamics on the pattern of linkage disequilibrium. Am J Hum Genet 68:198–207.

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES. 2001. Linkage disequilibrium in the human genome. Nature 411:199–204.

Santos-Lopes SS, Pereira RW, Wilson IJ, Pena SD. 2007. A worldwide phylogeography for the human X chromosome. PLoS One 2:e557.

Schroeder KB, Jakobsson M, Crawford MH, Schurr TG, Boca SM, Conrad DF, Tito RY, Osipova LP, Tarskaia LA, Zhadanov SI, Wall JD, Pritchard JK, Malhi RS, Smith DG, Rosenberg NA. 2009. Haplotypic background of a private allele at high frequency in the Americas. Mol Biol Evol 26:995–1016.

Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG. 2007. A private allele ubiquitous in the Americas. Biol Lett 3:218–223.

Service SK, DeYoung J, Karayiorgou M, Roos JL, Pretorious 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.

Service SK, Ophoff RA, Freimer NB. 2001. The genome-wide distribution of background linkage disequilibrium in a population isolate. Hum Mol Genet 10:545–551.

Slatkin M. 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet 9:477–485.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. Am J Hum Genet 76:449–462.

Stephens M, Smith N, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989.

Varilo T, Laan M, Hovatta I, Wiebe V, Terwilliger JD, Peltonen L. 2000. Linkage disequilibrium in isolated populations: Finland and a young sub-population of Kuusamo. Eur J Hum Genet 8:604–612.

Varilo T, Peltonen L. 2004. Isolates and their potential use in complex gene mapping efforts. Curr Opin Genet Dev 14:316–323.

Wang S, Bedoya G, Labuda D, Ruiz-Linares A. 2010. Brief communication: patterns of linkage disequilibrium and haplotype diversity at Xq13 in six Native American populations. Am J Phys Anthropol 142:476–480.

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.

Wang S, Ray N, Rojas W, Parra MV, Bedoya G, Gallo C, Poletti G, Mazzotti G, Hill K, Hurtado AM, Camrena B, Nicolini H, Klitz W, Barrantes R, Molina JA, Freimer NB, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Dipierri JE, Alfaro EL, Bailliet G, Bianchi NO, Llop E, Rothhammer F, Excoffier L, Ruiz-Linares A. 2008. Geographic patterns of genome admixture in Latin American Mestizos. PLoS Genet 4: e1000037.

Wilson DJ. 1999. Indigenous South Americans of the past and the present. An ecological perspective. Boulder: Westview Press.

Zarante I, Ossa D, Mendosa R, Valvuena G. 2000. Descripción etnográfica, demográfica y características en salud de las comunidades indígenas visitadas por La gran expedición humana. In: Vásquez AO, organizer. Geografía humana de Colombia: variación biológica y cultural en Colombia—Tomo 1. Santafé de Bogotá, DC: Editora Guadalupe. p 157–227.