Meta-Analysis of Brazilian Genetic Admixture and Comparison with Other Latin America Countries

RONALD RODRIGUES DE MOURA,[1,2]* ANTONIO VICTOR CAMPOS COELHO,[1,2] VALDIR DE QUEIROZ BALBINO,[1] SERGIO CROVELLA,[1,2] AND LUCAS ANDRÉ CAVALCANTI BRANDÃO[2,3]
[1]*Department of Genetics, Federal University of Pernambuco (UFPE), Recife, Pernambuco, Brazil*
[2]*Laboratory of Immunopatology Keizo Asami (LIKA/UFPE), Recife, Pernambuco, Brazil*
[3]*Department of Pathology, Federal University of Pernambuco (UFPE), Recife, Pernambuco, Brazil*

Objectives: This study aims at performing a systematic review and meta-analysis with the studies of genetic admixture inference of Brazilian population and to compare these results with the genetic admixture levels in other Latin American countries.

Methods: We searched for articles regarding the estimation of Brazilian genetic admixture published between 1980 and 2014 that used autosomal markers. Then, conducted meta-analyses at the whole-country and regional level. Finally, we compared the results of Brazil with other estimates from other South, Central and North American countries.

Results: We analyzed data from 25 studies in 38 different Brazilian populations. European (EUR) ancestry is the major contributor to the genetic background of Brazilians, followed by African (AFR), and Amerindian (AMR) ancestries. The pooled ancestry contributions were 0.62 EUR, 0.21 AFR, and 0.17AMR. The Southern region had a greater EUR contribution (0.77) than other regions. Individuals from the Northeast (NE) region had the highest AFR contribution (0.27) whereas individuals from the North regions had more AMR contribution (0.32). In the Latin America context, Brazil has the 5th high EUR contribution, the 12th for the AFR component and the 10th for the AMR ancestry.

Conclusions: Admixture proportions vary greatly among Brazilian populations and also through Latin America. More studies in the Center-West, North and NE regions are needed to capture a more complete picture of the genomic ancestry of Brazil. Am. J. Hum. Biol. 27:674–680, 2015.

Latin America populations exhibit varying degrees of genetic admixture due to different historical processes that have occurred since the end of the 15th century, leaving genetic traces of European (EUR), African (AFR), and Native American populations in the genomes of these individuals. Brazilian populations are not an exception to this general pattern (Salzano; Sans, 2014).

The colonization history of Brazil began in the 16th century, when the first Portuguese settlers (about a half million) started to mix with the indigenous populations (about 2.5 million) and then with AFR slaves (about 4 million) (IBGE, 2007). Moreover, after the establishment of the Republic of Brazil, in the 19th century, individuals from other nations migrated to Brazil (including Italians, Germans, and Japanese) (IBGE, 2007).

According to the last national census, Brazil has a population of about 200 million (IBGE, 2013). Genetic admixture has been directly influenced by this colonization process resulting in Brazil becoming a genetically trihybrid population. The genomic inheritance of EUR, AFR, and Amerindian (AMR) groups can be traced through the analysis of autosomal (Manta et al., 2013; Pena et al., 2011), sex chromosomes, and mitochondrial genetic information (Alves-silva et al., 2000; Palha et al., 2012).

Despite early insights about the Brazilian genetic ancestry emerging in the 1960s (Krieger et al., 1965), researchers have only extensively investigated the genetic contribution of EUR, AFR, and AMR ancestors to the genetic background of Brazil population from the 1980s onwards (Callegari-Jacques et al., 2003; Pena et al., 2009; Santos and Guerreiro, 1995, Schneider and Salzano, 1979).

Genomic admixture studies support the idea that it is not possible to use externally visible characteristics, such as hair, eye and skin color or even facial morphology, to infer the genetic ancestry of Brazilian individuals. These observations have clinical and social implications for affirmative policies implementation, case-control studies design, disease association studies and pharmacogenetics studies (Lins et al., 2011; Pena, 2005; Pena and Birchal, 2006; Suarez-kurtz et al., 2012).

To assess the ancestral proportion in Brazilian individuals, different parts of the genome have been analyzed, such as mitochondrial DNA (mtDNA) (Alves-silva et al. 2000), short tandem repeats (STR) (Callegari-Jacques et al., 2003), insertion/deletions (INDELS) (Santos et al., 2010) and single nucleotide polymorphisms (SNP) (Giolo et al., 2012). Criteria for ancestry informative markers (AIM) selection also varies among the studies, such as the absolute difference between the allele frequencies (δ), F statistics and the Informativeness for assignment (In) measure (Ding et al., 2011). Different methods to select the AIMs, the number of markers adopted in the study

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

Contract grant sponsor: Fundação de Amparo à Pesquisa do Estado de Pernambuco (FACEPE); Contract grant number: APQ-1530-2.02/12; Contract grant sponsor: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq); Contract grant number: 475350/2011-1.

*Correspondence to: Ronald Rodrigues de Moura, Department of Genetics, Federal University of Pernambuco (UFPE), Recife, Brazil.
E-mail: ronaldmoura1989@gmail.com

Accepted 11 February 2015

TABLE 1. Characterization of the studies included in the meta-analysis

Year	Author	State	Region	Type of marker	N° of markers	N° individuals	EUR[a]	AFR[a]	AMR[a]
1982	Schuller et al.	Amazonas	North	Blood systems	7	595	0.3600	0.1300	0.5100
1983	Santos et al.	Amazonas	North	Blood systems	3	954	0.6100	0.1200	0.2700
1984	Rosa et al.	Amazonas	North	Blood systems	8	811	0.4300	0.1400	0.4300
1987	Santos et al.	Pará	North	Blood systems	10	206	0.5700	0.1500	0.2800
1993	Guerreiro et al.	Pará	North	Blood systems	11	250	0.3800	0.1000	0.5200
1995	Ribeiro-dos-Santos et al.	Pará	North	Blood systems	13	500	0.4200	0.3300	0.2500
2005	Ferreira et al.	Maranhão	Northeast	STR/VNTR	5	177	0.4200	0.1900	0.3900
2006	Ferreira et al.	São Paulo	Southeast	STR	8	400	0.7900	0.1400	0.0700
2009	Scliar et al.	Minas Gerais	Southeast	STR/VNTR	13	234	0.6600	0.3200	0.0200
2010	Santos et al.	Pará	North	INDEL	48	196	0.6140	0.1170	0.2690
2010	Felix et al.	Bahia	Northeast	Alu/INDEL/ Restiction sites	7	289	0.4400	0.4900	0.0700
2010	Silva et al.	Minas Gerais	Southeast	SNP	14	24	0.5200	0.3900	0.0900
2010	Silva et al.	Minas Gerais	Southeast	SNP	14	30	0.7300	0.1900	0.0800
2011	Francez et al.	Amapá	North	STR	12	307	0.4600	0.1900	0.3500
2011	Martins et al.	São Paulo	Southeast	STR	15	403	0.7600	0.1800	0.0600
2011	Pena et al.	Pará	North	INDEL	40	203	0.7820	0.0770	0.1410
2011	Pena et al.	Ceará	Northeast	INDEL	40	82	0.7580	0.1330	0.1090
2011	Pena et al.	Bahia	Northeast	INDEL	40	147	0.6680	0.2440	0.0880
2011	Pena et al.	Rio de Janeiro	Southeast	INDEL	40	264	0.8610	0.0740	0.0650
2011	Pena et al.	Rio Grande do Sul	South	INDEL	40	189	0.8600	0.0500	0.0900
2011	Lins et al.	Distrito Federal	Centre-West	SNP/INDEL	13	189	0.6290	0.2540	0.1170
2011	Leite et al.	Distrito Federal	Centre-West	SNP	21	172	0.6900	0.2100	0.1000
2012	Francez et al.	Amapá	North	INDEL	48	130	0.5000	0.2900	0.2100
2012	Giolo et al.	São Paulo	Southeast	SNP	100	138	0.6100	0.2400	0.1500
2012	Pereira et al.	Pará	North	INDEL	46	226	0.5370	0.1680	0.2950
2012	Manta et al.	Rio de Janeiro	Southeast	INDEL	46	280	0.5520	0.3110	0.1370
2013	Cardena et al.	São Paulo	Southeast	INDEL	48	492	0.5740	0.2830	0.1430
2013	Manta et al.	Amazonas	North	INDEL	46	42	0.4590	0.1630	0.3780
2013	Manta et al.	Pernambuco	Northeast	INDEL	46	133	0.5680	0.2790	0.1530
2013	Manta et al.	Alagoas	Northeast	INDEL	46	104	0.5470	0.2660	0.1870
2013	Manta et al.	Mato Grosso do Sul	Centre-West	INDEL	46	84	0.5880	0.2590	0.1530
2013	Manta et al.	Minas Gerais	Southeast	INDEL	46	88	0.5920	0.2890	0.1190
2013	Manta et al.	Espírito Santo	Southeast	INDEL	46	92	0.7410	0.1340	0.1250
2013	Manta et al.	São Paulo	Southeast	INDEL	46	49	0.6290	0.2550	0.1160
2013	Manta et al.	Paraná	South	INDEL	46	21	0.7100	0.1750	0.1150
2013	Manta et al.	Santa Catarina	South	INDEL	46	20	0.7970	0.1140	0.0890
2013	Manta et al.	Rio Grande do Sul	South	INDEL	46	23	0.7290	0.1400	0.1300
2013	Queiroz et al.	Minas Gerais	Southeast	SNP	15	189	0.5030	0.3330	0.1640

[a]EUR, AFR, and AMR ancestries. STR: short tandem repeat; VNTR: variable number tandem repeat; INDEL: insertion/deletion; SNP: Single nucleotide polymorphism.

and different sampling strategies may contribute to divergent results among the studies (Manta et al., 2013).

In this context, we performed a systematic review with meta-analysis of the genetic admixture studies of the Brazilian population to estimate the pooled proportions of the EUR, AFR, and AMR contributions and compare these findings with the estimates of other countries in Latin America using the same approach.

MATERIAL AND METHODS

Literature search

The literature search was performed using the PubMed (www.ncbi.nlm.nih.gov/pubmed), Web of knowledge (www.webofknowledge.com), Scielo (www.scielo.org) and Google scholar (www.scholar.google.com) databases for studies published from 1980 to 2014 and also through the references cited in the articles. The key terms for the literature search were: "Brazilian Human Genetic admixture," "Brazilian Human ancestry," "Brazilian Human substructure," "Brazilian Human admixture" and "ancestry informative markers Brazilian population."

Inclusion/exclusion criteria

The inclusion criteria for the meta-analysis were (1) studies that provided the number of individuals studied per population (see the Supporting Information 2 for the Brazilian geographic distribution); (2) the use of nuclear markers to infer the genetic ancestry; and (3) explicitly information concerning the genetic proportions of the EUR, AFR, and AMR ancestries for each population (city) studied. In order to avoid possible bias, we also excluded from the meta-analysis studies focusing on completely or partially isolated populations, such as Asian and EUR colonies, indigenous tribes or AFR communities (*Quilombos*) remaining from the extensive period of slavery in the country (from 16th to late 18th century).

Data extraction

After excluding the articles that did not comply with the inclusion criteria, we retrieved the name of the first author, year of publication, type of genetic marker used (e.g., INDELs, SNP), number of genetic markers, geographic location of the population (e.g., São Paulo, Rio de Janeiro, Recife), sample size and the proportions of EUR, AFR, and AMR ancestries.

Statistical analysis

Standardized raw proportions of ancestry. The number of AIMs varied across the selected studies. To correct for any possible bias due to this variation, we developed two simple formulae that yielded raw proportions of EUR, AFR, and AMR ancestry:

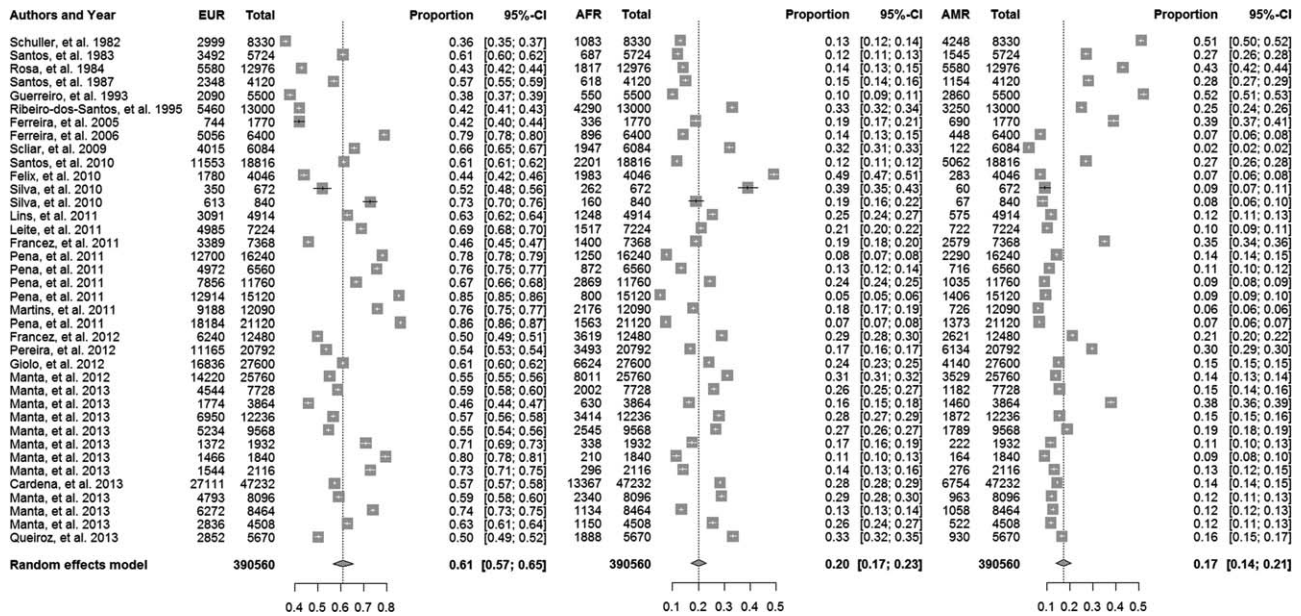

Fig. 1. Forest plot of the whole-Brazil EUR, AFR, and AMR ancestry proportions meta-analysis. The count of alleles, raw proportion, the exact confidence intervals (95% CI) and the weight (W) of each study included in the meta-analysis are shown. In the lower part of the plot, the pooled proportions with its 95% CI, and the values of the between-studies heterogeneity (τ^2 and I^2) can also be seen.

$$r = a_i x b_i x c_i x d \quad (1)$$

$$n = b_i x c_i x\ d \quad (2)$$

In the equations above, r is the probable number of alleles of a given ancestry; n is the total number of alleles sampled; a denotes a value between 0 and 1 which correspond to the inferred ancestry proportion for the study; b, represents the sample size; c, the number of AIMs used; and d represents the number of chromosome sets of the organism (humans are diploid). A raw proportion is simply a division between r and n (r/n).

As a demonstration, in a study with 196 individuals from Pará state, using 48 INDELs as AIMs, Santos et al. (2010) found values of 0.6140, 0.1170, and 0.2690 for the contributions of EUR, AFR, and AMR ancestries, respectively. Using the Eqs. (1) and (2) for the EUR estimate, we found $r_{eur} = 0.6140 \times 196 \times 48 \times 2$, which is 11,553; $n = 196 \times 48 \times 2$, which is 18,816. These numbers can represent the number of "European alleles" (r_{eur}) among all alleles (n). The same procedures can be done to estimate the AFR and AMR contribution.

With these formulae, we were able to obtain an ancestry estimate corrected by the number of AIMs used by each authors in their studies.

Meta-analysis

With the corrected ancestry estimates, we performed a meta-analysis for single proportions using the R package "meta" (R Core Team, 2013) for each ancestry separately. This package is able to calculate an overall proportion and its variance from datasets reporting a single proportion.

The I^2 and τ^2 measures were adopted to inform heterogeneity between-studies. Since high values of heterogeneity were found ($I^2 > 90\%$; $P < 0.1000$), we applied random effect model for all meta-analyses (DerSimonian and Laird, 1986). We calculated the transformed and pooled proportions using the Freeman-Tukey Double arcsine transformation (Freeman and Tukey, 1950), with its Clopper-Pearson confidence intervals, using both fixed and random effect models as well as the weight of the individual studies in the meta-analysis. All these results were displayed in the form of forest plots.

RESULTS

Characteristics of the elected studies

After a global literature search including papers published from 1980 to 2014, we selected 25 studies to be considered in our meta-analysis. In these studies, the nuclear genetic ancestry (EUR, AFR, and AMR) in 38 different populations from 17 different Brazilian states was assessed. Pará was the state with most populations studied (six) followed by São Paulo and Minas Gerais (both with five populations). Table 1 shows a summary of all studies included in the meta-analysis.

Regarding the type of AIMs used in these analyses, we subdivided the studies into four main categories: those using INDELs (Cardena et al., 2013; Kimura et al., 2013; Manta et al., 2012, 2013; Pena et al., 2011; Pereira et al., 2012; Santos et al., 2010), SNPs (Leite et al., 2011; Lins et al., 2011; Giolo et al., 2012; Queiroz et al., 2013; Silva et al., 2010), STR/VNTRs (Ferreira et al. 2005, 2006; Francez et al., 2011, 2012; Martins et al., 2011; Scliar et al., 2009) or blood group markers (Santos et al., 1983, 1987; Schuler et al., 1982; Rosa et al. 1984; Guerreiro et al. 1993; Ribeiro-dos-Santos et al., 1995). The only exception was the study of Felix et al. (2010), which employed a set of Alu elements, INDELs and restriction sites.

TABLE 2. *Pooled ancestry proportions by Brazilian's geopolitical region*

Region	Ancestry proportions (95% CI)		
	EUR[a]	AFR[a]	AMR[a]
North	0.51 (0.44–0.59)	0.16 (0.12–0.21)	0.32 (0.26–0.39)
Northeast	0.58 (0.48–0.66)	0.27 (0.19–0.34)	0.15 (0.10–0.21)
Centre-West	0.64 (0.58–0.69)	0.24 (0.21–0.27)	0.12 (0.09–0.16)
South	0.77 (0.69–0.85)	0.12 (0.06–0.19)	0.11 (0.09–0.12)
Southeast	0.67 (0.59–0.72)	0.23 (0.18–0.29)	0.10 (0.08–0.12)

[a]EUR, AFR, and AMR ancestries.

The first group employed sets of INDELs ranging from 40 to 48 mostly based on a larger dataset (Weber et al., 2002), sometimes sharing selected markers (Pena et al., 2011; Pereira et al., 2012; Santos et al., 2010) or the whole set in other cases (Manta et al., 2012, 2013; Pereira et al., 2012). The SNP group chose the AIMs from more sources (Bonilla et al., 2004; Fernández et al., 2003; Packer et al., 2006; Smith et al., 2004; Tian et al., 2006). They also share selected markers (Leite et al., 2011; Lins et al., 2011) or the whole set (Leite et al., 2011; Queiroz et al., 2013). Although the studies shared some variants in their sets, the source of the AIMs in the STR/VNTR group were not detailed with the exception of the study of Martins et al. (2011), which used a commercial kit. Finally, the blood group markers used in the six studies are detailed in the paper from Guerreiro et al. (1993).

The INDEL group elaborated its datasets using δ and F_{ST} measures. The SNP group was more divergent regarding the criteria of selection using combinations of δ, F_{ST}, In, and even principal component analysis (Giolo et al., 2012). The authors who used STR/VNTR and blood group markers did not specify the criteria used to select their AIMs.

Some studies used autosomal markers, but analyzed the data according to region and not according to state (Callegari-Jacques et al., 2003; Godinho et al., 2008; Lins et al., 2010) or they evaluated the ancestral proportions in semi or completely isolated populations (Callegari-jacques and Salzano, 1999; Maciel et al., 2011; Salzano et al., 1997). These studies were not included in the meta-analysis, although they served for comparison with our results.

Brazilian genetic admixture

We conducted the meta-analysis including the 25 studies that aggregated 3,90,560 alleles in 8,733 Brazilian individuals. The pooled proportions of European, AFR, and AMR ancestry with their respective confidence intervals are summarized in the forest plots (Fig. 1). The major contribution came from Europeans (0.61), followed by AFRs (0.20) and AMRs (0.17). It should be noted that the sum of the proportions is not equal to 1 (indeed, it is 0.98). This may happen because of rounding processes that occurred during the calculations. To circumvent this problem, we simply standardized the values dividing each pooled contribution by the sum of them. Therefore, for the EUR contribution we have 0.62 (i.e., 0.61/0.98), for AFR 0.21 (0.20/0.98) and 0.17 (0.17/0.98) for AMR ancestry.

We performed separated tests for the five Brazilian's geopolitical regions: North (N), NE, Centre-West (CW), South (S), Southeast (SE). The number of populations (and the total number of individuals in parenthesis) per region was as follows: N = 12 (4,420), NE = 6 (932), CW = 3 (445), S = 4 (253), and SE = 13 (2,683). Table 2 summarizes the pooled proportions of the EUR, AFR, and AMR ancestries per region.

The Southern region of Brazil had a greater EUR contribution (0.77) than other regions. The NE with 0.27 and the North with 0.32 were the regions with greater AFR and AMR contributions, respectively. The complete analysis of Brazilian genomic ancestry according to geographic region is described in the Supporting Information 1.

TABLE 3. *Genetic admixture proportions of EUR, AFR, and AMR parental populations in American countries*

Country	EUR	AFR	AMR	Reference
Peru	0.06	0.02	0.92	This study[a]
Ecuador	0.19	0.08	0.73	González-Andrade et al. (2007)
Mexico	0.31	0.06	0.62	This study[a]
Chile	0.42	0.02	0.56	Wang et al. (2008)
Guatemala	0.40	0.07	0.53	Wang et al. (2008)
Colombia	0.42	0.11	0.44	This study[a]
Argentina	0.54	0.03	0.42	This study[a]
Costa Rica	0.58	0.04	0.38	Ruiz-Narváez et al. (2010)
WC-USA[b]	0.56	0.08	0.36	Halder et al. (2009)
Venezuela	0.60	0.14	0.25	This study[a]
Brazil	0.62	0.21	0.17	This study[a]
EC-USA[b]	0.65	0.18	0.17	Halder et al. (2009)
Dominica	0.28	0.56	0.16	Torres et al. (2013)
Puerto Rico	0.65	0.20	0.14	This study[a]
Nicaragua	0.69	0.20	0.11	Nuñez et al. (2010)
Uruguay	0.84	0.06	0.10	Hidalgo et al. (2005)
Trinidad and Tobago	0.16	0.75	0.09	Torres et al. (2013)
Jamaica	0.10	0.82	0.08	Torres et al. (2013)
St. Lucia	0.18	0.75	0.07	Torres et al. (2013)
Grenada	0.12	0.81	0.07	Torres et al. (2013)
St. Thomas	0.17	0.77	0.06	Torres et al. (2013)
St. Vincent	0.13	0.81	0.06	Torres et al. (2013)
St. Kitts and Nevis	0.08	0.86	0.06	Torres et al. (2013)
AA-USA[b]	0.16	0.81	0.04	Halder et al. (2009)
EA-USA[b]	0.98	0.01	0.01	Halder et al. (2009)
Cuba	0.73	0.26	0.01	Diaz-Horta et al. (2010)
Bahamas	0.04	0.96	0.00	Simms et al. (2010)
Haiti	0.04	0.96	0.00	Simms et al. (2010)

The data are listed in descendent order of AMR contribution.
[a]With the exception of the references used in the Brazilian meta-analyses, the complete references for this table can be found in the Supporting Information 3.
[b]AA-USA stands for AFR American from United States of America (USA); EA is European American; EC is East Coast Hispanics; WC is West Coast Hispanics.

Admixture proportions in Latin America countries

Recently, Salzano and Sans (2014) published a review which discussed the genetic admixture in Latin American populations, although they did not explore the data using a meta-analytical approach. Therefore, based on their article, we expanded our analysis in order to verify the proportions of parental populations (AFR, AMR, and EUR) in other countries from Latin America using meta-analyses of genetic admixture studies carried out in populations from Argentina (20 populations), Colombia (25), Mexico (23), Peru (25), Puerto Rico (8), and Venezuela (5) using the same inclusion criteria applied for Brazilian meta-analysis. These results are summarized in Table 3. In this table, we also listed admixture proportions for other American countries' populations (including USA), which did not have sufficient published data to allow meta-analysis to better represent genetic admixture in Latin America.

After conducting the meta-analyses, the Mexican population has a pooled 0.31, 0.06, and 0.62 of EUR, AFR, and

AMR ancestry proportions, respectively. In Central America, the Nicaraguan population had the highest EUR and AFR contributions (0.69 and 0.20, respectively), whilst the highest AMR contribution was described in Guatemala (0.53). Among the South American countries, Peruvians showed 0.92 of AMR contribution, whereas Brazilians had 0.21 AFR contribution. Uruguayans showed the highest EUR contribution (0.84).

Among Caribbean Islands, Haiti, and the Bahamas had almost total AFR contributions (0.96), while Cuba had 0.73 of EUR ancestry and Dominica 0.16 AMR ancestry.

DISCUSSION

In the last two decades, genetic admixture in Brazilian populations has been a matter of concerted investigation. These investigations attempted to delineate the composition of Brazilians' genetic background in uniparental (Alves-Silva et al., 2000; Carvalho-silva et al., 2001) and biparental contexts (Santos and Guerreiro, 1995; Schneider and Salzano, 1979; Salzano et al., 1997; Callegari-jacques et al., 2003; Parra et al., 2003).

Although several studies have demonstrated the genetic admixture in various populations through all Brazil and other Latin America countries, there is no systematic review compiling these studies in order to establish overall results based on available data. This was the main objective of the present work.

Throughout our search, we found a considerable number of articles that used only uniparental markers, such as Marrero et al. (2005), Guerreiro-Junior et al. (2009), and Bernardo et al. (2014) who used mtDNA, and Silva et al. (2006) and Carvalho et al. (2010) who used Y-chromosome markers. However, the majority of data came from Southeastern and Southern regions, which will bias the meta-analysis results. Therefore, we decided to concentrate on only autosomal markers for this article.

Considering all Brazilian populations studied, we found that the pooled EUR, AFR and AMR ancestry proportions were 0.62, 0.21, and 0.17, respectively. These results are in agreement with other findings describing the highest contribution of EURs, followed by AFRs and AMRs, to Brazilians (Godinho et al., 2008; Lins et al., 2010). We observed high values of between-study heterogeneity. This heterogeneity might be due to the differences in sample sizes and different number and sets of markers used. For example, Manta and et al. (2013) and Scliar et al. (2009), using INDELs and microsatellite data, respectively, found different ancestry proportions for Brazilians (see Table 1). We can also hypothesize that different contexts of admixture processes occurred in the study populations as well as social and cultural influences that may account for this heterogeneity.

In the Latin American context, Brazil has the 5th highest EUR genetic ancestry (0.62) after Uruguay (0.84), Cuba (0.73), Nicaragua (0.69) and Puerto Rico (0.65). In general, EUR ancestors made a larger genetic contribution in the Atlantic side of the American continent, whereas the AMR contribution occurred predominantly at the Pacific side. Peru, Ecuador and Mexico are the three countries with the highest American Native contributions (0.92, 0.73 and 0.62, respectively). These countries were formerly the core of Inca (Ecuador and Peru) and Aztec (Mexico) Empires. Despite the massive depopulation due to epidemics, exploitation and war during the Spanish invasion, natives from these populations were more involved in the admixture process than other indigenous populations (Salzano and Callegari-Jacques, 1988). Brazil is the 10th population with respect to AMR proportion (0.17).

Although Brazil has the highest AFR contribution among South American countries (0.21), it is the 12th for AFR contribution if we consider Latin America as a whole, since Caribbean countries have the highest estimates for AFR ancestry, such as Bahamas and Haiti with 96% of AFR contribution. During the colonization of the Caribbean Islands by British, France, Spanish and other EUR countries, almost all indigenous people were killed or deported, requiring them to import a large number of AFR Slaves to work on sugar-cane and coffee plantations (Knight, 1997). These slaves played a vital role in the historical and genetic composition of Caribbean countries. In Haiti, for example, after the Haitian revolution in the early 19th century, much of the French colonizers left the island (Pamphile, 2001).

The same scenario evidenced at national scale has also been observed at regional scale in Brazil. The EUR contribution is the highest in all five regions of Brazil. With the exception of Lins et al. (2010), who found more AMR contribution in Central-West than in North Brazil, our results agree with other authors who reported an increased value for the AMR ancestry in the North in comparison with the other regions and an higher EUR contribution in the South (Callegari-jacques et al., 2003; Godinho et al., 2008; Manta et al., 2013; Pena et al., 2011).

Although the Southern region of Brazil has the lowest AFR proportion (0.12), this value is high compared to the bordering countries such as Argentina (0.03) and Uruguay (0.06). On the other hand, in the Northern region the AMR proportion is the highest (0.32) in the country, also higher when compared to Venezuela (0.25) but lower if compared to Colombia (0.44) and Peru (0.92).

The regional genomic distribution found in Brazilian is linked with the different colonization history of each region. For example, the South region received successive migration cycles of EURs during the 16th (Portuguese) and 19th (Germans and Italians) centuries. Moreover, the majority of AFR slaves arrived and settled in Brazil's NE and SE before moving to other Brazilian regions (Conrad, 1973; Levy, 1974; IBGE, 2007). These observations help to explain why, according to our meta-analysis, the South region has 77% of EUR genomic ancestry, whereas NE has 58%. The opposite can be seen when we look at the AFR contribution: 12% in the South and 27% in NE.

Using F-statistics, two studies reported significant genetic distance between populations from different regions from Brazil. Lins et al. (2010) obtained significant values when comparing the South region population with those from other regions. In the other study, Manta et al. (2013) also verified divergence between urban populations from South and populations from North, NE and CW regions. It is important to bear in mind that, although these authors reported significant distance between Brazilian populations, the magnitude of these differences did not reach more than 10%, which represents a low to moderate genetic divergence according to Wright's qualitative guidelines (Wright, 1978). If compared with proxies of parental populations (e.g. CEU and YRI populations from HapMap Project), the F_{ST} can reach values of 44%

depending on the proportion of a given ancestry present in the Brazilian population that is been compared (Manta et al., 2013).

Some articles not included in our meta-analysis evaluated admixture in AFR (*Quilombos*) and AMR communities. These studies found AFR contributions ranging from 0.32 to 0.92 (Kimura et al., 2013; Maciel et al., 2011; Scliar et al., 2009) and AMR contributions ranging from 0.25 to 0.97 (Callegari-jacques and Salzano, 1999; Manta et al., 2013; Salzano et al., 1997). These values are superior to the pooled AFR and AMR contribution of our meta-analysis for the whole of Brazil (0.21 and 0.17, respectively).

Some of the studies included in our meta-analysis investigated the correlation between self-reported skin color and genomic ancestry in Brazilian population (Leite et al., 2011; Lins et al., 2011; Pena et al., 2011; Queiroz et al., 2013). From those papers, Leite et al. (2011) used quantitative measures to evaluate the associations between melanin index, self-reported skin color and genomic ancestry among pairs of siblings.

Although they found statistical differences between self-reported skin color and melanin index, there was considerable overlap between groups. Furthermore, the correlations between self-reported skin color *versus* genomic ancestry and melanin index *versus* genomic ancestry also had considerable overlap between skin color categories and continental ancestry. These results are in agreement with other studies in Latin America that pointed to similar findings (Parra et al., 2004; Ruiz-Linares et al., 2014).

Moreover, self-perception of ancestry is biased not only by skin color but also by other phenotypic traits, such as iris and hair color, which tend to overestimate EUR ancestry, and hair type and some facial characteristics that may overestimate AFR ancestry. Apart from phenotypic traits, socioeconomic factors also contribute to this self-perception, with wealth and education also tending to overestimate EUR ancestry (Ruiz-Linares et al., 2014).

Considering the geographic distribution of the study populations (Supporting Information 2), 10 states had no information about the nuclear genetic ancestry of their populations; most of them are from the North and NE regions. Moreover, only three studies were conducted in populations from CW. Therefore, further studies are required to better elucidate and correctly describe the genomic ancestry of CW, North and NE Brazilian regions. It is well worth noting that the study of the Brazilian ancestry is also making an important contribution to clinical research, providing fundamental information about the development of genetic association studies searching for disease-related markers or clinical epidemiological analyses (Lins et al., 2011; Pinto et al., 2012; Suarez-kurtz et al., 2012): Moreover, ancestry studies are subsidizing others issues, such as sociological debates (Pena, 2005; Pena and Birchal, 2006).

CONCLUSION

In the present work, we used a meta-analytic approach to compile various studies on the genetic ancestry of Brazilian populations, based on nuclear markers, and compared the results with the admixture data for other countries in Latin America. We concluded that the pooled proportions of EUR, AFR, and AMR ancestries in Brazil (globally considered) are 0.62, 0.21, and 0.17, respectively. These values, mainly for AFR and AMR contributions, are intermediate when compared to other Latin American countries. At the regional level, as expected, the highest AMR contribution occurred in the Northern region, the highest AFR contribution in the Northeastern region and the highest EUR contribution is in the Southern region. More studies in the CW, North and NE regions are needed to capture the whole landscape of the genomic ancestry of Brazil.

LITERATURE CITED

Alves-silva J, Santos MS, Guimarães PEM, Ferreira ACS, Bandelt HJ, Pena SDJ, Prado VF. 2000. The ancestry of Brazilian mtDNA lineages. Am J Hum Genet 67:444–461.

Bernardo S, Hermida R, Desidério M, Silva DA, Carvalho EF. 2014. mtDNA ancestry of Rio de Janeiro population, Brazil. Mol Biol Rep 41: 1945–1950.

Bonilla C, Shriver MD, Parra EJ, Jones A, Fernández JR. 2004. Ancestral proportions and their association with skin pigmentation and bone mineral density in Puerto Rican women from New York City. Hum Genet 115:57–68.

Callegari-jacques SM, Grattapaglia D, Salzano FM, Salamoni SP, Grossetti SG, Ferreira ME, Hutz MH. 2003. Historical genetics: spatio-temporal analysis of the formation of the Brazilian population. Am J Hum Biol 15:824–834.

Callegari-jacques SM, Salzano FM. 1999. Brazilian Indian/non- Indian interactions and their effects. Cienc Cult 51:166–174.

Cardena MMSG, Ribeiro-dos-Santos AKC, Santos S, Mansur AJ, Pereira AC, Fridman C. 2013. Assessment of the relationship between self-declared ethnicity, mitochondrial haplogroups and genomic ancestry in Brazilian individuals. PloS One 8:e62005.

Carvalho M, Brito P, Lopes V, Andrade L, Anjos MJ, Real FC, Gusmão L. 2010. Analysis of paternal lineages in Brazilian and African populations. Genet Mol Biol 33:422–427.

Carvalho-silva DR, Santos FR, Rocha J, Pena SDJ. 2001. The Phylogeography of Brazilian Y-Chromosome Lineages. Am J Hum Genet 68:281–286.

Conrad R. 1973. The destruction of Brazilian slavery 1850-1888. Oakland, CA: University of California. p 368.

DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Contr Clin Trials 7:177–188.

Ding L, Wiener H, Abebe T, Altaye M, Go RCP, Kercsmar C, Grabowski G, Martin LJ, Hershey GKK, Chakorborty R, Baye TM. 2011. Comparison of measures of marker informativeness for ancestry and admixture mapping. BMC Genomics 12:622–640.

Felix GES, Abe-sandes K, Bonfim TM, Bendicho MT, Guedes R, Brandão CJF, Torres AJL, Brites C, Netto EM, Meyer R, Freire SM. 2010. Ancestry informative markers and complete blood count parameters in Brazilian blood donors. Rev Bras Hematol Hemoter 32:282–285.

Fernández JR, Shriver MD, Beasley TM, Rafla-demetrious N, Parra E, Albu J, Nicklas B, Ryan AS, McKeigue PM, Hoggart CL, Weinster RL, Allison DB. 2003. Association of African genetic admixture with resting metabolic rate and obesity among women. Obesity Res 11:904–911.

Ferreira FL, Leal-mesquita ER, Santos SEB, Ribeiro-dos-Santos AKC. 2005. Genetic characterization of the population of São Luís, MA, Brazil. Genet Mol Biol 28:22–31.

Ferreira LB, Mendes-junior CT, Wiezel CEV, Luizon MR, Simões AL. 2006. Genomic Ancestry of a Sample Population from the State of São Paulo, Brazil. Am J Hum Biol 18:702–705.

Francez PAC, Ribeiro-Rodrigues EM, Frazão GF, Borges NDR, Santos SEB. 2011. Allelic frequencies and statistical data obtained from 12 codis STR loci in an admixed population of the Brazilian Amazon. Genet Mol Biol 34:35–39.

Francez PAC, Ribeiro-Rodrigues EM, Santos SEB. 2012. Allelic frequencies and statistical data obtained from 48 AIM INDEL loci in an admixed population from the Brazilian Amazon. Forensic Sci Int 6:132–135.

Freeman MF, Tukey JW. 1950. Transformations Related to the Angular and the Square Root. Ann Math Stat 21:607–611.

Giolo SR, Soler JMP, Greenway SC, Almeida MAA, Andrade M, Seidman JG, Seidman CE, Krieger JE, Pereira AC. 2012. Brazilian urban population genetic structure reveals a high degree of admixture. Eur J Hum Genet 20:111–116.

Godinho NMO, Contijo CC, Diniz MECG, Falcão-Alencar G, Dalton GC, Amorim CEG, Barcelos RSS, Klautau-Guimarães MN, Oliveira, SF. 2008. Regional patterns of genetic admixture in South America. Forensic Sci Int 1:329–330.

Gravlee CC, Dressler WW, Bernard HR. 2005. Skin color, social classification, and blood pressure in southeastern Puerto Rico. Am J Public Health 95:2191–2197.

Guerreiro JF, Ribeiro-dos-Santos AK, Santos EJM, Cayres IMV, Santos SEB. 1993. Genetic Structure and Demography of the Human Population of Óbidos, in the Brazilian Amazon. Braz J Genet 16:1075–1084.

Guerreiro-Junior V, Bisso-Machado R, Marrero A, Hünemeier T, Salzano FM, Bortolini MC. 2009. Genetic signatures of parental contribution in black and white populations in Brazil. Genet Mol Biol 32:1–11.
Knight FW. 1997. General history of the Caribbean—volume III: the slave societies of the Caribbean. London, UK: UNESCO Publishing. p 380.
Instituto Brasileiro de Estatística (IBGE). 2007. Brasil: 500 anos de povoamento. Rio de Janeiro: IBGE. p 237.
Instituto Brasileiro de Estatística (IBGE). 2013. Atlas do Censo Demográfico 2010. Rio de Janeiro: IBGE. p 216.
Kimura L, Ribeiro-rodrigues EM, Auricchio MTBM, Vicente JP, Santos SEB, Mingroni-Netto RC. 2013. Genomic Ancestry of Rural African-Derived Populations from Southeastern Brazil. Am J Hum Biol 25:35–41.
Krieger H, Morton NE, Azevedo E, Freire-Maia N, Yasuda N. 1965. Racial admixture in Northeastern Brazil. Ann Hum Genet 29:113–125.
Leite TKM, Fonseca RMC, França NM, Parra EJ, Pereira RW. 2011. Genomic Ancestry, Self-Reported "'Color'" and Quantitative Measures of Skin Pigmentation in Brazilian Admixed Siblings. PloS One 6:e27162.
Levy MSF. 1974. The role of international migration on the evolution of the Brazilian population (1872 to 1972). Rev saude publ 8:49–90.
Lins TC, Vieira RG, Abreu BS, Gentil P, Moreno-lima R, Oliveira RJ, Pereira RW. 2011. Genetic Heterogeneity of Self-Reported Ancestry Groups in an Admixed Brazilian Population. J Epidemiol 21:240–245.
Lins TC, Vieira RG, Abreu BS, Grattapaglia D, Pereira RW. 2010. Genetic Composition of Brazilian Population Samples Based on a Set of Twenty Eight Ancestry Informative SNPs. Am J Hum Biol 22:187–192.
Lins TC, Vieira RG, Grattapaglia D, Pereira RW. 2011. Population analysis of vitamin D receptor polymorphisms and the role of genetic ancestry in an admixed population. Genet Mol Biol 34:377–385.
Maciel LGL, Ribeiro-rodrigues EM, Santos NPC, Ribeiro-dos-Santos AKC, Guerreiro JF, Santos S. 2011. Afro-Derived Amazonian Populations: inferring Continental Ancestry and Population Substructure. Hum Biol 83:627–636.
Manta FSN, Pereira R, Vianna R, Araújo ARB, Gitaí GDL, Silva DA, Wolfgramm EV, Pontes IMP, Aguiar JI, Moraes MO, Carvalho EF, Gusmão L. 2013. Revisiting the genetic ancestry of Brazilians using autossomal AIM-INDELs. PloS One 8:e75145.
Manta FSN, Pereira R, Caiafa A, Silva DA, Gusmão L, Carvalho EF. 2012. Analysis of genetic ancestry in the admixed Brazilian population from Rio de Janeiro using 46 autosomal ancestry-informative indel markers. Ann Hum Biol 40:1–5.
Marrero AR, Leite FP, Carvalho BA, Peres LM, Kommers TC, Cruz IM, Bortolini MC. 2005. Heterogeneity of the Genome Ancestry of Individuals Classified as White in the State of Rio Grande do Sul, Brazil. Am J Hum Biol 17:496–506.
Martins JA, Figueiredo RF, Yoshizaki CS, Paneto GG, Cicarelli RMB. 2011. Genetic data of 15 autosomal STR loci: an analysis of the Araraquara population colonization (São Paulo, Brazil). Mol Biol Rep 38: 5397–5403.
Packer BR, Yeager M, Burdett L, Welch R, Beerman M, Qi L, Sicotte H, Staats B, Acharya M, Crenshaw A, Eckert A, Puri V, Gerhard DS, Chanock SJ. 2006. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. Nucleic Acids Res 34:D617–D621.
Palha T, Ribeiro-rodrigues E, Ribeiro-dos-Santos AKC, Santos S. 2012. Fourteen short tandem repeat loci Y chromosome haplotypes: genetic analysis in populations from northern Brazil. Forensic Sci Int 6:413–418.
Pamphile LD. 2001. Haitians and African Americans. Gainseville, FL: University Press of Florida.
Parra FC, Amado RC, Lambertucci R, Rocha J, Antunes CM, Pena SDJ. 2003. Color and genomic ancestry in Brazilians. PNAS 100:177–182.
Parra EJ, Kittles RA, Shriver MD. 2004. Implications of correlations between skin color and genetic ancestry for biomedical research. Nat Genet 36:S54–S60.
Pena SDJ. 2005. Razões para banir o conceito de raça da medicina brasileira. Hist Cienc Saude 12:319–346.
Pena SDJ, Bastos-rodrigues L, Pimenta JR, Bydlowski SP. 2009. DNA tests probe the genomic ancestry of Brazilians. Braz J Med Biol Res 42: 870–876.
Pena SDJ, Birchal TS. 2006. A inexistência biológica vesus a existêcia social de raças humanas: pode a ciência instruir o etos social? Rev USP 68:10–21.
Pena SDJ, Pietro GD, Fuchshuber-moraes M, Genro JP, Hutz MH, Kehdy FSG, Kohlrausch F, Magno LAV, Montenegro RC, Moraes MO, Moraes MEA, Moraes MR, Ojopi EB, Perini JA, Racciopi C, Ribeiro-dos-Santos AKC, Rios-Santos F, Romano-Silva MA, Sortica VA, Suarez-kurtz, G. 2011. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. PloS One 6:e17063.
Pereira R, Phillips C, Santos C, Santos SEB, Carracedo A, Gusmão L. 2012. Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing. PloS One 7:e29684.
Pinto P, Salgado CG, Santos N, Alencar DO, Santos S, Hutz MH, Ribeiro-dos-Santos AKC. 2012. Polymorphisms in the CYP2E1 and GSTM1 genes as possible protection factors for leprosy patients. PloS One 7: e47498.
Queiroz EM, Santos AM, Castro IM, Machado-Coelho GL, Cândido APC, Leite TM, Pereira RW, Freitas RN. 2013. Genetic composition of a Brazilian population: the footprint of the gold cycle. Genet Mol Res 12: 5124–5133.
R Core Team. 2013. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/
Ribeiro-dos-Santos ÁKC, Santos EJM, Guerreiro JF, Santos SEB. 1995. Demographic and genetic structure of the population of Castanhal, in the Amazon region of Brazil. Braz J Genet 18:469–474.
Rosa VL, Salzano FM, Franco MH, Freitas MJ. 1984. Blood genetic studies in five Amazonian populations. Braz J Genet 7:569–582.
Ruiz-Linares A, Adhikari K, Acuña-Alonzo V, Quinto-Sanchez M, Jaramillo C, Arias W, Gonzalez-José R. 2014. Admixture in Latin America: geographic structure, phenotypic diversity and self-perception of ancestry based on 7,342 individuals. PLoS Genet 10: e1004572.
Salzano FM and Callegari-Jacques SM. 1988. South American Indians: a case study in evolution. Oxford, UK: Clarendon Press. p 259.
Salzano FM, Callegari-jacques SM, Weimer TA, Franco MHLP, Hutz MH, Petzl-erler ML. 1997. Electrophoretic protein polymorphisms in Kaingang and Guarani Indians of Southern Brazil. Am J Hum Biol 9:502–512.
Salzano FM, Sans M. 2014. Interethnic admixture and the evolution of Latin American populations. Genet Mol Biol 37:151–170.
Santos SEB, Salzano FM, Helena M, Franco LP, Freitas MJM. 1983. Mobility, genetic markers, susceptibility to malaria and race mixture in Manaus, Brazil. J Hum Evol 12:373–381.
Santos S, Guerreiro JF, Salzano FM, Weimer TA, Hutz MH, Franco MHLP. 1987. Mobility, blood genetic traits and race mixture in the Amazonian population of Oriximiná. Braz J Genet 10:745–759.
Santos S, Guerreiro J. 1995. The indigenous contribution to the formation of the population of the Brazilian Amazon Region. Rev Bras Genet 18: 311–315.
Santos NPC, Ribeiro-rodrigues EM, Ribeiro-dos-Santos AKC, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, Santos SEB. 2010. Assessing individual interethnic admixture and population substructure using a 48–insertion-deletion (INSEL) ancestry-informative marker (AIM) panel. Hum Mutat 31:184–190.
Schneider H, Salzano FM. 1979. Gm allotypes and racial admixture in two Brazilian populations. Hum Genet 53:101–105.
Schüler L, Salzano FM, Franco MHLP, Freitas MJM, Mestriner MA, Simões AL. 1982. Demographic and blood genetic characteristics in an Amazon population. J Hum Evol 11:549–558.
Scliar MO, Vaintraub MT, Vaintraub PMV, Fonseca CG. 2009. Admixture analysis with forensic microsatellites in Minas Gerais, Brazil: the ongoing evolution of the capital and of an African-derived community. Am J Phys Anthropol 139:591–595.
Silva DA, Carvalho EF, Costa G, Tavares L, Amorim A, Gusmão L. 2006. Y-chromosome genetic variation in Rio de Janeiro population. Am J Hum Biol 18:829–837.
Silva MCF, Zuccherato LW, Soares-Souza GB, Vieira ZM, Cabrera L, Herrera P, Balqui J, Romero C, Jahuira H, Gilman RH, Martins ML, Tarazona-santos E. 2010. Development of two multiplex minisequencing panels of ancestry informative SNPs for studies in Latin Americans: an application to populations of the State of Minas Gerais (Brazil). Genet Mol Res 9:2069–2085.
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, Thé G, Essex M, Sankalé J, Moore JH, Poku K, Phair JP, Goedert JJ, Vlahov D, Williams SM, Tishkoff SA, Winkler CA, De la Vega FM, Woodgate 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.
Suarez-kurtz G, Pena SDJ, Struchiner CJ, Hutz MH. 2012. Pharmacogenomic diversity among Brazilians: influence of ancestry, self-reported color, and geographical origin. Front Pharmaco 3:1–7.
Tian C, Hinds DA, Shigeta R, Kittles R, Ballinger DG, Seldin MF. 2006. Genome-wide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. Am J Hum Genet 79:640–649.
Weber JL, David D, Heil J, Fan Y, Zhao C, Marth G. 2002. Human diallelic insertion/deletion polymorphisms. Am J Hum Genet 71:854–862.
Wright S. 1978. Variability within and among natural populations. Chicago, IL: University of Chicago, v. 4, 590 p.