

are larger than those previously observed on the basis of classical markers (Batista dos Santos et al. 1999; Santos et al. 1999).

mtDNA analysis has been used extensively during the past 10 years, since the pioneering works of Vigilant (1990), Stoneking et al. (1991), and Vigilant et al. (1991). Phylogeographic analysis of mtDNA lineages from all over the world has led to the identification of mtDNA haplogroups that are specific to either Africans, Europeans, or Asians/Amerindians (Torroni et al. 1993, 1994, 1996, 1998; Chen et al. 1995; Richards et al. 1996; Watson et al. 1997; Kivisild et al. 1999a, 1999b; Macaulay et al. 1999; Metspalu et al. 1999). Haplogroup allocation of a given mtDNA lineage allows the assessment of its (sub)continental origin, so that the matrilineal ancestry of admixed populations can be evaluated well (Torroni et al. 1995; Bravi et al. 1997; Green et al. 2000; Rando et al. 1999).

In the present article, we follow this approach by sequencing part of the control region and by screening specific RFLP sites, to better understand the extent of the matrilineal genetic contribution of Europeans, Africans, Amerindians, and Asians to the gene pool of present-day Brazilians.

Subjects and Methods

Samples

We analyzed 247 unrelated Brazilian individuals (mainly classified as “white” in Brazil and belonging to the middle and upper-middle classes) who came from four of the five major geographic regions of the country (fig. 1). According to the Instituto Brasileiro de Geografia Estatística, responsible for the census in Brazil, 51.6% of Brazilians in 1996 classified themselves as white. In detail, there were 99 individuals from the southeastern (mostly from the state of Minas Gerais), 50 from the southern (states of Rio Grande do Sul, Santa Catarina, and Paraná), 48 from the northern (states of Amazonas, Pará, Rondônia, and Acre), and 50 from the northeastern (state of Pernambuco) regions. Thirty-seven individuals were students or staff in our laboratory, whereas 210 were randomly chosen unrelated participants in paternity-testing studies. Written consent was obtained from all participants, and all analyses were performed anonymously.

mtDNA Control-Region Amplification and Sequencing

The nucleotide sequence of mtDNA hypervariable segment I (HVS-I), between nucleotide positions (np) 16060 and 16362, d

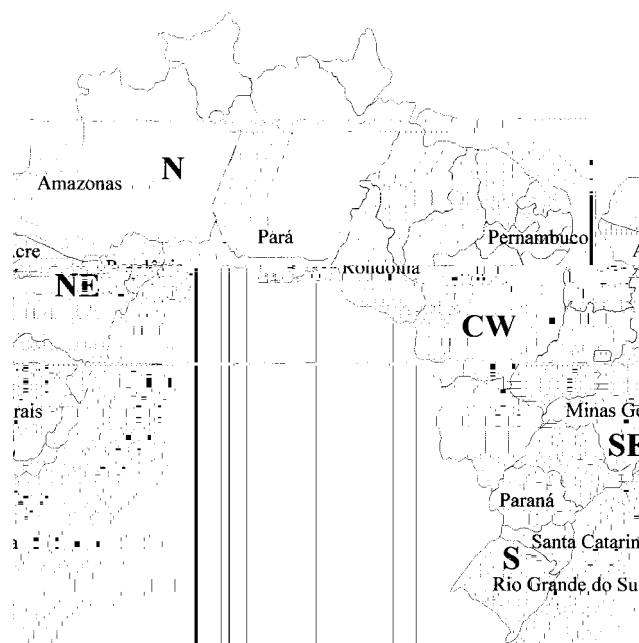


Figure 1 Five major geographic regions of Brazil: N = northern, NE = northeastern, SE = southeastern, S = southern, and CW = central west. Brazilian states from which the mtDNA lineages of the present study have mainly been sampled are indicated by name.

Table 1
RFLP Polymorphisms Used to Identify mtDNA Haplogroups and Geographic Origin

HAPLOGROUP	CHARACTERISTIC RESTRICTION SITE(S)	STATUS ^a		
		Sub-Saharan African	Native American	European
L1a	+3592 <i>Hpa</i> I, +11641 <i>Hae</i> III	+	–	–
L1b	+3592 <i>Hpa</i> I, –7055 <i>Alu</i> I, +2349 <i>Mbo</i> I	+	–	–
L1c	+3592 <i>Hpa</i> I, +9070 <i>Taq</i> I, +12810 <i>Rsa</i> I	+	–	–
L2	+3592 <i>Hpa</i> I, +16389 <i>Hin</i> FI	+	–	–
L3b	+10084 <i>Taq</i> I	+	–	–
L3d	–3592 <i>Hpa</i> I, –8616 <i>Mbo</i> I	+	–	–
L3e	–3592 <i>Hpa</i> I, +2349 <i>Mbo</i> I	+	–	–
A	+663 <i>Hae</i> III	–	+	–
B	9-bp deletion	–	+	–
C	–13259 <i>Hinc</i> II	–	+	–
D	–5176 <i>Alu</i> I	–	+	–
H	–7025 <i>Alu</i> I	–	–	+
V	–4577 <i>Nla</i> III	–	–	+
HV	–14766 <i>Mse</i> I	–	–	+
U	+12308 <i>Hin</i> FI	–	–	+
K	–9052 <i>Hae</i> II	–	–	+
J	–13704 <i>Bst</i> NI	–	–	+
T	+13366 <i>Bam</i> HI, +15606 <i>Alu</i> I	–	–	+
I	–4529 <i>Hae</i> II, +8249 <i>Ava</i> II, +16389 <i>Bam</i> HI, +10032 <i>Alu</i> I	–	–	+
W	+8249 <i>Ava</i> II, –8994 <i>Hae</i> III	–	–	+
X	–1715 <i>Dde</i> I	–	+	+

^a A plus sign (+) denotes that the haplogroup is indig89 75 0 TEs;ind127.9lussign (

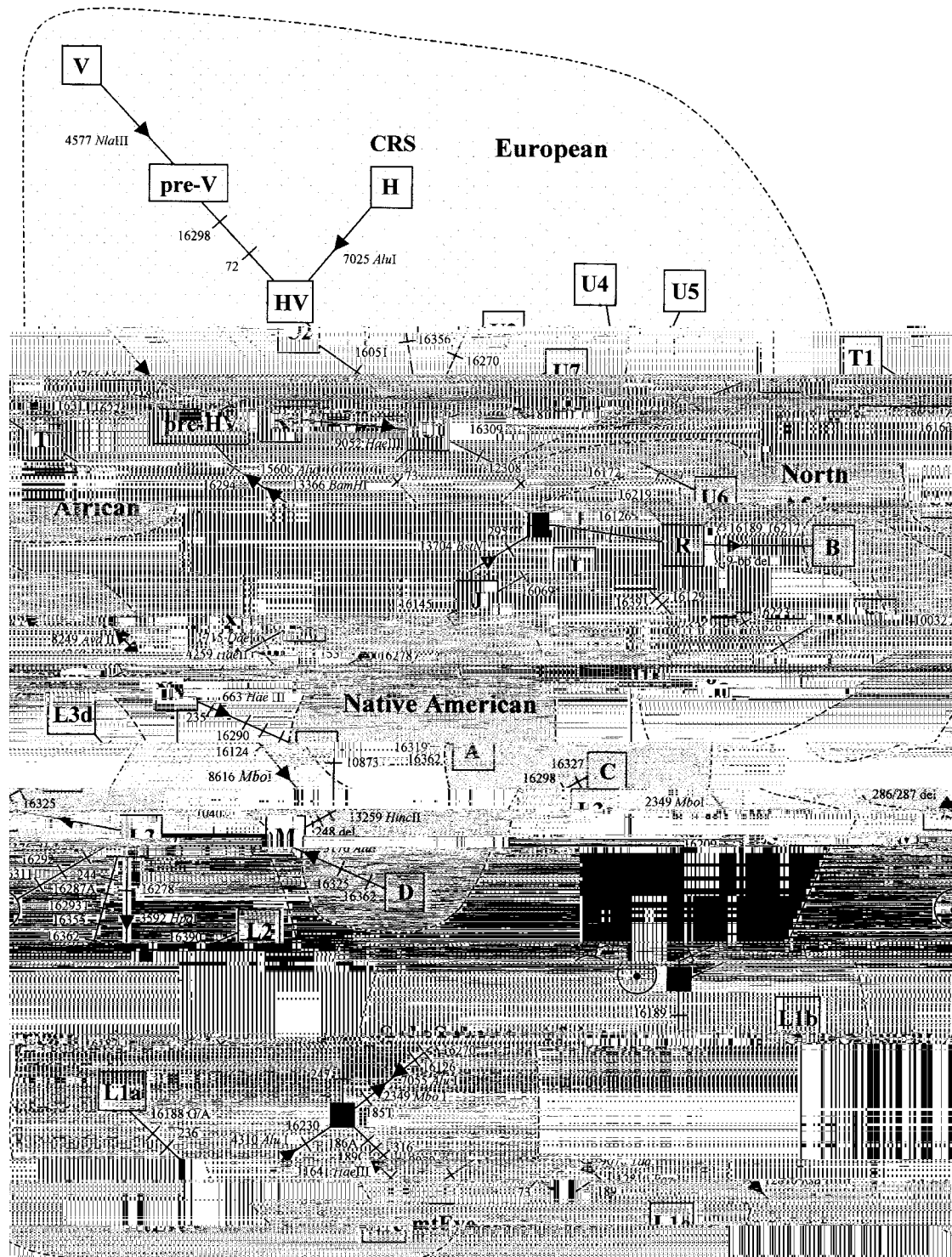


Figure 2 Classification tree highlighting selected diagnostic sites and positions for haplogroups present in the Brazilian sample (see tables 1 and table 6). Each square represents the root node of the respective haplogroup, with the acronym inscribed; two central/eastern-African haplogroups, represented by circles, are only partially characterized (T. Kisivild, personal communication). “CRS” indicates the revised reference sequence (Andrews et al. 1999). Numbers along links refer to RFLP sites (with arrows pointing to presence of sites) or transitions, unless a single-letter suffix indicates a transversion. Note that some diagnostic sites and positions, especially in the control region, have undergone

HVS-I Haplotypes and Their Regional Distribution in Brazil

448

BR41	1	T.....	T.....G.....	C.....CT.....	C
BR42	1	A.....	T.....A.....	C.....CT.....T.	C
BR43	1	A.....	T.....G.....	C.....CT.....	C
BR44	2	A.....	T.....G.....	C.....CT.....	C
BR45	2	A.....	T.....G.....	C.....CT.....	C
BR46	1	T.....	T.....T.....	C.....C.....	D
BR47	1	T.....	T.....T.....	C.....C.....	D
BR48	1	A.....	T.....T.....	C.....C.....	D
BR49	1	A.....	T.....T.....	C.....C.....	D
BR50	1	C.....	T.....T.....	C.....C.....	D
BR51	1	C.....	T.....T.....	C.....C.....	D
BR52	1	T.....	T.....G.....	C.....C.....	D
BR53	1	A.....	T.....G.....	C.....C.....	D
African:					
BR54	1	T.....	T.....G.....	C.....T.....	L1a
BR55	2	A.....	T.....G.....	C.....T.....	L1a
BR56	1	A.....	T.....G.....	C.....T.....	L1a
BR57	1	A.....	T.....G.....	C.....T.....	L1a
BR58	1	C.....	T.....G.....	C.....T.....	L1a
BR59	1	C.....	T.....G.....	C.....T.....	L1b
BR60	1	T.....	T.....T.....	C.....C.....	L1b
BR61	1	C.....	T.....T.....	C.....C.....	L1b
BR62	1	A.....	T.....T.....	C.....C.....	L1c
BR63	1	C.....	T.....T.....	C.....C.....	L1c
BR64	1	T.....	T.....T.....	C.....C.....	L1c
BR65	1	A.....	T.....T.....	C.....C.....	L1c
BR66	2	A.....	T.....T.....	C.....C.....	L1c
BR67	1	T.....	T.....T.....	C.....C.....	L1c
BR68	1	C.....	T.....T.....	C.....C.....	L1c
BR69	1	A.....	T.....T.....	C.....C.....	L1c
BR70	1	A.....	T.....T.....	C.....C.....	L1c
BR71	1	C.....	T.....T.....	C.....C.....	L1c
BR72	1	A.....	T.....T.....	C.....C.....	L1c
BR73	1	C.....	T.....T.....	C.....C.....	L1c
BR74	1	T.....	T.....T.....	C.....C.....	L2
BR75	1	C.....	T.....T.....	C.....C.....	L2
BR76	1	CT.....	T.....T.....	C.....C.....	L2
BR77	1	T.....	T.....T.....	C.....C.....	L2
BR78	1	T.....	T.....T.....	C.....C.....	L2
BR79	1	A.....	T.....T.....	C.....C.....	L2
BR80	1	A.....	T.....T.....	C.....C.....	L2
BR81	2	T.....	T.....T.....	C.....C.....	L2
BR82	1	A.....	T.....T.....	C.....C.....	L2
BR83	1	A.....	T.....T.....	C.....C.....	L2
BR84	1	A.....	T.....T.....	C.....C.....	L2
BR85	1	C.....	T.....T.....	C.....C.....	L2
BR86	1	A.....	T.....T.....	C.....C.....	L2
BR87	1	C.....	T.....T.....	C.....C.....	L2
BR88	1	C.....	T.....T.....	C.....C.....	L3d
BR89	1	C.....	T.....T.....	C.....C.....	L3d
BR90	1	C.....	T.....T.....	C.....C.....	L3d

BR131	1	1	1	T.....	C..C.....	pre*V
BR132			1		C.....	pre*V
BR133	1			T.....	C.....	V
BR134	1			C.....	C.....	V
BR135		1		A.....	C.....	V
BR136		1		C.....	C.....	V
BR137	2				C.....T.....	V
BR138		1		A.....	HV*	
BR139	1		1	C.....	U2	
BR140	1		1	T.....	U4	
BR141	1			C.....	T..A.....C.....	U5b*
BR142		1		C.....	T.....	U5b*
BR143	1		1	T.C.....	U5b1	
BR144		1			G..T.....	U7
BR145	1				C.....	K
BR146		1	1	C.....	C.....	K
BR147			1	A.....	C.....	K
BR148	1			C.....	C.....	K
BR149			1	C.....	C..A.....	K
BR150		1		C.....	C..A.....	K
BR151	1	2	1	T.....	pre*HV	
BR152	1		2	C.....	J*	
BR153		1		T.....	J*	
BR154		1	1	C.....	T.....	J*
BR155	1			C..A.....	T.....T.....	J1*
BR156		2		C.....	T.....	J1b1
BR157			1	C.....	T..T..C.....	T*
BR158	1	1	1	C.....	T..T..C.....	T*
BR159		1		C..T.....	T.TC..C.....	T*
BR160	1			C.....	T..T.....	T*
BR161	1			C.....	T..T..C.....	T*
BR162		1		C.....A.....	T..T.....	T*
BR163		1		C.....A.....	T.....T.....	T*
BR164		1		T.C.....A.....	T.....T.....	T*
BR165	1			C.....	T..T.....	T*
BR166		1	1	C.....G.....	T.....	T1
BR167		1		A.....	T.....	I
BR168			1	C.....	T.....	X
BR169	1			C.....	G..T.....	X
BR170	1			C.....	T.....	X
Total	99	50	50	48		

^a BR13 and BR14 share the loss of 3534 *Ddel*, whereas controls BR4, BR5, BR15, and BR16 were +3534 *Ddel*. BR132 has C at np 72, and for BR139 the U2-characteristic transition at np 16051 has been confirmed.

^b SE = southeastern, S = southern, NE = northeastern, N = northern.

^c The prefix "16" has been deleted from all (three-digit) numbers. Length083polymorphisms)TheN

Table 4
Frequency of Continent-Specific mtDNA Haplotypes in the Brazilian mtDNA Pool

CONTINENTAL FRACTION	FREQUENCY				
	Brazil	Northern	Northeastern	Southeastern	Southern
Native American	.33	.54	.22	.33 ^a	.22
African	.28	.15	.44	.34	.12
European	.39	.31	.34	.31	.66

^a Excludes the single lineage of confirmed Asian ancestry.

the Gulf of Guinea (central Africa), however, have an

Table 5
Haplogroup Frequencies within the Three Continental Fractions of Brazilian mtDNA Pool

HAPLOGROUP	F
------------	---

regions of the country. The highest Amerindian influence was observed in populations of the Amazonian region, where a study analyzing 11 urban populations by use of nuclear markers observed an average of 41% Amerindian ancestry (Santos and Guerreiro 1995). Recent mtDNA analysis of another population from the northern region showed an even higher Amerindian contribution (59% [Batista dos Santos et al. 1999]). In our sample, we also observed a high Amerindian influence in the northern region (54%), corroborating the mtDNA data obtained by Batista dos Santos et al. (1999). In the other regions of Brazil, the genetic contribution from Amerindians is also markedly higher for mtDNA than for nuclear DNA: 22% (table 4) versus 13% (Krieger et al. 1965; Franco et al. 1982; Conceição et al. 1987) in the northeastern region and 22% (table 4) versus 11% (Dornelles et al. 1999) in the southern region. For the southeastern region, where we have detected 33% frequency of mtDNA lineages of Amerindian ancestry, not a single study with nuclear markers has yet been performed, probably because one did not anticipate measurable Amerindian genetic influence on urban populations (Salzano 1997). As for African admixture in the white Brazilian population, the picture is similar to what we have seen for the Amerindian genetic input: mtDNA analysis (table 4) suggests a higher contribution than that by nuclear markers, for which 12% in the northern region (Santos et al. 1996a), 36% in the northeastern region (Franco et al. 1982; Arpini-Sampaio et al. 1999), and 10% in the southern region (Dornelles et al. 1999) were reported; again, no nuclear data are available for the southeastern region.

The allocation of haplogroups to continents (as indicated in table 1) is, of course, not absolutely clear-cut. For instance, the European haplogroups H and U5 do occur in the sub-Saharan mtDNA pool, albeit in only two founder types (bearing transitions at np 16145 and 16222 for the H type and transitions at np 16189, 16192, 16270, and 16320 for the U5 type), possibly transmitted by Berbers or, even earlier, during the Saharan Neolithic age (Rando et al. 1999). None of these particular mtDNA lineages occur in our Brazilian sample. Similarly, northern-African U6 haplotypes have penetrated the Sahara and are found sporadically from the west (Senegal) to the east (Kenya). We consider it most plausible that the four U6 lineages in our sample have come from western Africa. On the other hand, African haplotypes were also transmitted, in low numbers, to Europe, especially to the Mediterranean area. African mtDNA lineages, then, constitute erratic outliers in the respective mtDNA samples, for instance, such as the L1c lineage in the British data of Piercy et al. (1993).

There is one caveat with regard to the distinction between European mtDNA haplotypes and Native

American ones: haplogroup X is shared by western Eurasia and North America (Brown et al. 1998; Smith et al. 1999), although there is as yet no compelling evidence for the occurrence of haplogroup X in Central or South America. The three X haplotypes that we detected in the Brazilian sample are certainly of European ancestry, since BR169 and BR170 do not bear the np-200 transition that is characteristic of (most of) the Native American haplogroup X (Brown et al. 1998), whereas BR168 (for which no HVS-II information is available) bears a transition (namely, at np 16248) already observed in Europe (Richards et al. 1998).

The distinction between Asian and Native American mtDNA haplotypes is more intricate inasmuch as haplogroups A–D are of Asian origin. Fortunately, the Native American A, C, and D founder HVS-I and HVS-II types can be distinguished from Asian haplotypes by mutations that are virtually absent, or at least rare, in Asia. The transition at np 16325 is (almost) diagnostic for Native American C and D haplotypes; the 2-bp deletion in HVS-II seems to be characteristic of Native American C (table 6; also see Ginther et al. 1993; Kolman and Bermingham 1997), since it has not been reported in Asian mtDNAs so far (Lee et al. 1997). The “Beringian” transition at np 16111 is seen in most Native American A lineages but is virtually absent in Asia (Horai and Hayasaka 1990; Horai et al. 1993; Torroni et al. 1993; Kolman et al. 1996; Lee et al. 1997). Thus, only haplotype BR16, which, incidentally, matches an mtDNA lineage from Hokkaido (Horai et al. 1996),

HVS-II Haplotypes in Southeastern Brazil

HVS-II Haplotypes in Southeastern Brazil

[illegible]

^b Haplogroup-diagnostic nucleotide positions are underlined.

able for Africa or most parts of Europe—in particular, Italy. Considering that 30% of the European immigrants (including the Portuguese colonizers) to Brazil came from Italy (Salzano and Freire Maia 1967; Callegari-Jacques and Salzano 1999), one can expect that a considerable number of mtDNA lineages in the Brazilian sample have Italian ancestry. I-28206tr54*[anBra-crmsts) ItaBra-

In principle, it should be possible to narrow the matrilineal ancestry of Brazilians to a geographic scale narrower than that of the (sub)continents (e.g., see Alves-Silva et al. 1999). This, in general, requires extensive phylogeographic studies of the populations from the potential source areas, which, at present, are not avail-

of Brazil. It needs to be emphasized that genetic distances, although trivial to compute, between the Brazilian mtDNA sample and mtDNA samples representing potential source populations would not allow the calculation of reliable admixture proportions, as demonstrated by Rando et al. (1999) in the case of the mixed population from the Canary Islands.

One could also use a reverse approach and infer the mtDNA profile of a source population from that of the target mixed population (given sufficient information on the other participating source populations). For instance, no mtDNA data for Angola are available—yet, since Angola was the major source of African slaves brought to Brazil, we can make inferences on how the mtDNA pool of Angolans would look: we should expect (i) a considerable number of L3e lineages—in particular, those bearing the np-16327 transition, also observed in the Herero and in other southern African populations

- Bortolini MC, Weimer TA, Franco MH, Salzano FM, Layrisse Z, Schneider H, Schneider MP, et al (1992) Genetic studies in three South American black populations. *Gene Geogr* 6: 1–16
- Bortolini MC, Weimer TDA, Salzano FM, Callegari-Jacques SM, Schneider H, Layrisse Z, Bonatto SL (1995) Evolutionary relationships between black South American and African populations. *Hum Biol* 67:547–559
- Bortolini MC, Zago MA, Salzano FM, Silva-Junior WA, Bonatto SL, da Silva MC, Weimer TA (1997b) Evolutionary and anthropological implications of mitochondrial DNA variation in African Brazilian populations. *Hum Biol* 69:141–159
- Bravi CM, Sans M, Bailliet G, Martinez-Marignac VL, Portas M, Barreto I, Bonilla C, et al (1997) Characterization of mitochondrial DNA and Y-chromosome haplotypes in a Uruguayan population of African ancestry. *Hum Biol* 69: 641–652
- Brown MD, Hosseini SH, Torroni A, Bandelt H-J, Allen JC, Schurr TG, Scozzari R, et al (1998) mtDNA haplogroup X: an ancient link between Europe/western Asia and North America? *Am J Hum Genet* 63:1852–1861
- Callegari-Jacques SM, Salzano FM (1999) Brazilian Indian/non-Indian interactions and their effects. *Ciênc Cult* 51: 166–174
- Chen YS, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet* 57:133–149
- Conceição MM, Salzano FM, Franco MHL, Weimer TA, Krieger H (1987) Demography, genetics, and race admixture in Aracaju, Brazil. *Braz J Genet* 10:313–331
- Côrte-Real HB, Macaulay VA, Richards MB, Hariti G, Issad MS, Cambon-Thomsen A, Papiha S, et al (1996) Genetic diversity in the Iberian Peninsula determined from mitochondrial sequence analysis. *Ann Hum Genet* 60:331–350
- Curtin PD (1969) *The Atlantic slave trade: a census*. University of Wisconsin Press, Madison
- Dornelles CL, Callegari-Jacques SM, Robinson WM, Weimer TA, Franco MHL, Hickmann AC, Geiger CJ, et al (1999) Genetics, surnames, grandparents' nationalities, and ethnic admixture in Southern Brazil: do the patterns of variation coincide? *Genet Mol Biol* 22:151–161
- Easton RD, Merriwether DA, Crews DE, Ferrell RE (1996) mtDNA variation in the Yanomami: evidence for additional New World founding lineages. *Am J Hum Genet* 59:213–225
- Forster P, Harding R, Torroni A, Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59:935–945
- Francalacci P, Bertranpetit J, Calafell F, Underhill PA (1996) Sequence diversity of the control region of mitochondrial DNA in Tuscany and its implications for the peopling of Europe. *Am J Phys Anthropol* 100:443–460
- Franco MH, Weimer TA, Salzano FM (1982) Blood polymorphisms and racial admixture in two Brazilian populations. *Am J Phys Anthropol* 58:127–132
- Ginther C, Corach D, Penacino GA, Rey JA, Carnese FR, Hutz MH, Anderson A, et al (1993) Genetic variation among the Mapuche Indians from the Patagonian region of Argentina: mitochondrial DNA sequence variation and allele frequencies of several nuclear genes. *EXS* 67:211–219
- Graven L, Passarino G, Semino O, Boursot P, Santachiara-Benerecetti S, Langaney A, Excoffier L (1995) Evolutionary correlation between control region sequence and restriction polymorphisms in the mitochondrial genome of a large Senegalese Mandenka sample. *Mol Biol Evol* 12:334–345
- Green LD, Derr JN, Knight A (2000) mtDNA affinities of the peoples of north-central Mexico. *Am J Hum Genet* 66: 989–998
- Guerreiro JF, Ribeiro-dos-Santos AKC, Santos EJM, Vallinoto ACR, Cayres-Vallinoto IMV, Aguiar GFS, Santos SEB (1999) Genetical-demographic data from two Amazonian populations composed of descendants of African slaves: Pacoval and Curiau. *Genet Mol Biol* 22:163–167
- Helgason A, Sigurðardóttir S, Gulcher JR, Ward R, Stefánsson K (2000) mtDNA and the origin of the Icelanders: deciphering signals of recent population history. *Am J Hum Genet* 66:999–1016
- Horai S, Hayasaka K (1990) Intraspecific nucleotide sequence differences in the major noncoding region of human mitochondrial DNA. *Am J Hum Genet* 46:828–842
- Horai S, Hayasaka K, Kondo R, Tsugane K, Takahata N (1995) Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc Natl Acad Sci USA* 92:532–536
- Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S, Tajima K (1993) Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol Biol Evol* 10:23–47
- Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, et al (1996) mtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan. *Am J Hum Genet* 59:579–590
- Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, et al (1999a) Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Curr Biol* 9:1331–1334
- Kivisild T, Kaldma K, Metspalu M, Parik J, Papiha S, Villems R (1999b) The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the old world. In: Papiha G, Deka R, Chakra-

- Lee SD, Shin CH, Kim KB, Lee YS, Lee JB (1997) Sequence variation of mitochondrial DNA control region in Koreans. *Forensic Sci Int* 87:99–116
- Macaulay V, Richards M, Hickey E, Vega E, Cruciani F, Guida V, Scozzari R, et al (1999) The emerging tree of West Eurasian mtDNAs: a synthesis of control-region sequences and RFLPs. *Am J Hum Genet* 64:232–249
- Mateu E, Comas D, Calafell F, Perez-Lezaun A, Abade A, Bertranpetit J (1997) A tale of two islands: population history and mitochondrial DNA sequence variation of Bioko and São Tomé, Gulf of Guinea. *Ann Hum Genet* 61:507–518

- Native American, and European mitochondrial DNAs in Cubans from Pinar del Rio Province and implications for the recent epidemic neuropathy in Cuba. *Hum Mutat* 5: 310–317
- Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, et al (1996) Classification of European mtDNAs from an analysis of three European populations. *Genetics* 144:1835–1850
- Torroni A, Lott MT, Cabell MF, Chen YS, Laverigne L, Wallace DC (1994) mtDNA and the origin of Caucasians: identification of ancient Caucasian-specific haplogroups, one of which is prone to a recurrent somatic duplication in the D-loop region. *Am J Hum Genet* 55:760–776
- Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, Smith DG, et al (1993) Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet* 53:563–590
- Torroni A, Schurr TG, Yang CC, Szathmary EJ, Williams RC,