AMERICAN JOURNAL OF HUMAN BIOLOGY 16:57–67 (2004)

Original Research Article

Binary and Microsatellite Polymorphisms of the Y-Chromosome in the Mbenzele Pygmies From the Central African Republic

VALENTINA COIA,[1] ALESSANDRA CAGLIÀ,[2] BARBARA ARREDI,[2,3] FRANCESCO DONATI,[1] FABRÍCIO R. SANTOS,[3] ARPITA PANDYA,[3] LUCA TAGLIOLI,[4] GIORGIO PAOLI,[4] VINCENZO PASCALI,[2] GIOVANNI DESTRO-BISOL,[1]* AND CHRIS TYLER-SMITH[5]
[1]*Department of Animal and Human Biology, University "La Sapienza," Rome, Italy*
[2]*Istituto di Medicina Legale, Catholic University, Rome, Italy*
[3]*Department of Biochemistry, University of Oxford, Oxford, UK*
[4]*Department of Ethology, Ecology and Evolution, University of Pisa, Pisa, Italy*
[5]*The Wellcome Trust Sanger Institute, Hinxton, UK*

ABSTRACT This study analyzes the variation of six binary polymorphisms and six microsatellites in the Mbenzele Pygmies from the Central African Republic. Five different haplogroups (B2b, E(xE3a), E3a, P and BR(xB2b,DE,P)) were observed, with frequencies ranging from 0.022 (haplogroup P) to 0.609 (haplogroup E3a). A comparison of haplogroup frequencies indicates a close genetic affinity between the Mbenzele and the Biaka Pygmies, a finding consistent with the common origin and the geographical proximity of the two populations. The haplogroups P, BR(xB2b,DE,P) and E(xE3a), which are rare in sub-Saharan Africa but common in western Eurasia, were observed with frequencies ranging from 0.022 (haplogroup P) to 0.087 (haplogroup E(xE3a)). Thirty different microsatellite haplotypes were detected, with frequencies ranging from 0.022 to 0.152. The Mbenzele share the highest percent of microsatellite haplotypes with the Biaka Pygmies. Five out seven haplotypes which are shared by the Mbenzele and Biaka Pygmies belong to haplogroup E3a, which suggests that they are of Bantu origin. The plot based on F_{st} genetic distances calculated using microsatellite data provides a picture of population relationships which is in part congruent and in part complementary to that obtained using haplogroup frequencies. Finally, the Mbenzele and Biaka Pygmies were found to be markedly more genetically similar using Y-chromosomal than autosomal microsatellites. We suggest that this could be due to the higher phylogenetic stability of Y-chromosome and to the effect of the male-biased gene flow during the Bantu expansion. Am. J. Hum. Biol. 16:57–67, 2004.

In recent years the interest of genetic anthropologists in the Y-chromosome has increased noticeably and its use in population studies has become comparable to that of mitochondrial DNA (mtDNA). By virtue of its paternal inheritance and lack of recombination along most of its length, the Y-chromosome constitutes the primary source of information for the reconstruction of paternal lineages of human populations. Besides its usefulness as a natural counterpart to mtDNA, the Y-chromosome has another feature which is relevant in anthropological studies, namely, the presence of polymorphisms with different evolutionary rates. In fact, the intense scrutiny of Y-chromosome variation has led to the identification of numerous slow mutating and phylogenetically stable polymorphisms, such as single nucleotide polymorphisms (SNPs), and fast-mutating polymorphisms, such as microsatellites. In the past, Y-chromosomal polymorphisms have proved to be particularly useful in a number of anthropological and genetic surveys. Such studies have investigated different time-scales, with the origin of anatomically modern humans (Hammer, 1997, 1998, 2001) and the study of Jefferson and Cohen lineages at the two extremities (Foster et al., 1998; Thomas et al., 1998). The use of SNPs

Contract grant sponsors: University of Rome "La Sapienza"; M.I.U.R. (cofinanziamento 2003–2005).

*Correspondence to: Giovanni Destro-Bisol, Department of Animal and Human Biology, University "La Sapienza", P.le A. Moro 5, 00185 Roma, Italia. E-mail: destrobisol@uniroma1.it

Received 10 April 2003; Revision received 22 September 2003; Accepted 26 September 2003

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/ajhb.10236

also in combination with microsatellites has proved to be a very valuable approach for understanding complex population histories. In the case of African populations, Thomas et al. (2000) identified Y-chromosomes of probable Bantu and Semitic origin among the Lemba, the so-called "Black Jews of Southern Africa," using this composite approach. Considering the African Pygmies specifically, Cruciani et al. (2002) estimated that their proportion of Y-chromosomes of probable Bantu origin is very high, larger than 50%.

As part of a long-term project regarding molecular variation in Central Africa (Cameroon and Central African Republic; see http://www.scienzemfn.uniroma1.it/labantro), we analyzed six microsatellites and six SNPs in the Mbenzele Pygmies from the Central African Republic. The Mbenzele Pygmies were analyzed for several loci (see below), but not for Y-chromosomal markers. Therefore, with this study we have achieved a more complete genetic characterization of this population. After the description of Y-chromosomal variation in the Mbenzele Pygmies, we examine the relationships between variation of SNP and microsatellite polymorphisms. In order to provide a comparative framework, we use unpublished Y-chromosomal SNP data for the Bamileke and Ewondo from Cameroon which were analyzed for the same loci studied in the Mbenzele Pygmies. Finally, SNP and microsatellite data are used to analyze the genetic differentiation of the Mbenzele Pygmies and some other sub-Saharan populations, with a view to evaluating the validity of microsatellite data in inferring population relationships.

THE POPULATION

The Mbenzele belong to the western cluster of African Pygmies (Murdock, 1959) which includes Pygmies from the Central African Republic, Congo Brazza, and Southern Cameroon. Like the Biaka Pygmies studied by Cavalli-Sforza (1986a,b), this population belongs to the Aka subgroup. At present, there are about 2,000 Mbenzele, mostly settled in the southwestern region of the Central African Republic (Corvin 1988). Blood samples were collected in the village of Mbelemboke (3.20N; 16.10E), which was created by Catholic missionaries in 1975. Previous articles concerning genetic variation in the Mbenzele Pygmies include studies of the ApoB 3′HVR minisatellite locus (Destro-Bisol et al., 2000a), of 10 autosomal microsatellite loci (Destro-Bisol et al., 2000b), of 10 protein loci (Coia et al., 2002). Most of the analyses use unpublished data relative to the Bamileke and Ewondo from Cameroon (see Spedini et al., 1999, for further details on these populations). These two samples are distinct from those studied by Cruciani et al. (2002) and were analyzed for comparison purposes. The readers should refer to Cruciani et al. (2002) for high-resolution Y-chromosome haplogrouping and to Caglià et al. (2003) for a more complete study of microsatellite haplotypes in the Bamileke and Ewondo.

MATERIALS AND METHODS

A total of 120 apparently healthy and unrelated males was analyzed. Forty-six blood samples were collected from as many Mbenzele individuals in the village of Mbelemboke (3.20N; 16.10E), in the southwestern region of the Central African Republic. The Bamileke (49 individuals) and Ewondo (25 individuals) blood samples were collected in the Bandjoun (5°21′N, 10°30′E) and Febe villages (3°54′N, 11°30′E), respectively. Informed consent was obtained from all donors.

Laboratory analyses

Specimens collected in K_3EDTA were maintained at 4°C (for < 7 days) until their arrival at the laboratory of Anthropology of the University of Rome "La Sapienza." Aliquots of 0.5 ml each were frozen and then thawed and the red cells selectively lysed by $1 \times$ SSC. White cells were pelleted, the DNA extracted by a standard protocol, and quantified by direct comparison with standards on agarose minigels. We analyzed six biallelic polymorphisms or their phylogenetic equivalents (SRY-1532; 92R7 (or M45); 50f2(P) (or M112); YAP (or M145); SRY-8299 (or M96); sY81) identifying the haplogroups reported in Table 1 which were named according to the YCC (Y-Chromosome Consortium, 2002) and Jobling et al. (1997). Typing of PCR products was performed by restriction enzyme digestion or primer extension and mass spectrometry (Paracchini et al., 2002). Six microsatellite loci (DYS19, DYS389I, DYS390, DYS391, DYS392, and DYS393) located in the nonrecombining portion of the Y-chromosome (Roewer et al., 1996) were also typed.

TABLE 1. Allelic status of Y-chromosome haplogroups observed in the present study[a,b]

	Haplogroup (HG)[a]				
Nomenclature	P 1	BR(xB2b,DE,P) 2	B2b 6	E3a 8	E(xE3a) 21
Locus					
YAP, M145	− (0)[b]	− (0)	− (0)	+ (1)	+ (1)
SRY-8299,M96	G (0)	G (0)	G (0)	A (1)	A (1)
SY81, M2[c]	A (0)	A (0)	A (0)	G (1)	A (0)
50F2(P), M112	G (0)	G (0)	C (1)	G (0)	G (0)
92R7, M45	T (1)	C (0)	C (0)	C (0)	C (0)
SRY-1532[d]	A (0)	G (1)	G (1)	G (1)	G (1)

[a]Haplogroup nomenclature is in accordance with the YCC (2002) (upper) or Jobling et al. (1997) (lower).
[b]The notation (0) and (1) indicates the ancestral and the derived status, respectively.
[c]Also referred to as DYS271.
[d]Also referred to as SRY10831.

Data analyses

The Arlequin software (Schneider et al., 1997) was used to compute the following parameters: 1) haplotype diversity (h) according to equation 8.7 in Nei (1987); 2) the mean number of pairwise differences (π); 3) values of F_{st} (Weir and Cockerham, 1984; Michalakis and Excoffier, 1996; Reynolds et al., 1983) and R_{st} (Slatkin, 1995) genetic distances and their statistical significance; 4) The same software was used to perform the analysis of molecular variance (AMOVA). To visualize the genetic relationships among the groups examined, we built principal component plots using haplogroup frequencies and genetic distance matrices, using the Statistica software (release 5.0).

RESULTS AND DISCUSSION

Y-chromosomal variation in the Mbenzele Pygmies

Haplogroups. All individuals were classified into haplogroups B2b, E(xE3a), E3a, P, and BR(xB2b,DE,P). Haplogroup frequencies of the Mbenzele Pygmies ranged from 0.022 (haplogroup P) to 0.609 (haplogroup E3a) (Table 2). Haplogroup E3a, which carries the M2 mutation (also referred to as DYS271), has been associated with the expansion of Bantu-speaking populations from central western towards southern Africa (Hammer et al., 1997; Scozzari et al., 1999; Underhill et al., 2001). Its high frequency in the Mbenzele confirms that this is the most frequent haplogroup in most sub-Saharan populations. Among broad groups of populations, its frequency ranges from 0.86 in East Africans to 0.20 in Khoisan, but it accounts for 50% or more of the chromosomes in all groups except the Khoisan (Hammer et al., 1998, where our haplogroup E3a corresponds closely to their haplotype 5). The estimated frequency of haplogroup E3a among the Mbenzele Pygmies is similar to that of the Biaka (0.650; Cruciani et al., 2002), which suggests that the Bantu contribution to the male-specific gene pool of these two Western Pygmy populations is comparable. The relatively high frequency (0.239) of haplogroup B2b is a characteristic feature of Mbenzele. It has been found also in Khoisan (0.078; Cruciani et al., 2002), Biaka Pygmies (0.300; Cruciani et al., 2002), and Eastern Pygmies (0.250; Cruciani et al., 2002). The haplogroups P and BR (xB2b,DE,P) were found in one and two individuals, respectively, while haplogroup E(xE3a) was found in four individuals. All of these haplogroups are rare in sub-Saharan Africa but common in western Eurasia, especially Europe, and may represent admixture mediated by Bantu farmers. Haplogroup A, characterized by an A at SRY-1532, was not found among the Mbenzele or other Pygmies, although the corresponding haplotypes (2 + 1A) make up over 50% of the Khoisan (Hammer et al., 1998), thus demonstrating a clear difference between these two groups.

TABLE 2. Frequencies of Y-chromosome haplogroups in the Mbenzele Pygmies

Haplogroup	Frequency
P	0.022
BR(xB2b,DE,P)	0.043
B2b	0.239
E3a	0.609
E(xE3a)	0.087

TABLE 3. Compound Y-chromosome haplotypes observed among the Mbenzele Pygmies

Compound haplotype	Microsatellite haplotype[a]						HG[b]	N (% frequency)
	DYS19	DYS389I	DYS390	DYS391	DYS392	DYS393		
MBE01	14	12	24	11	11	13	E(xE3a)	1 (0.022)
MBE02	14	13	21	10	11	14	P	1 (0.022)
MBE03	14	13	24	10	11	13	E(xE3a)	1 (0.022)
MBE04	14	13	24	11	11	13	E(xE3a)	2 (0.043)
MBE05	15	11	24	11	11	14	B2b	1 (0.022)
MBE06	15	11	25	11	11	14	B2b	1 (0.022)
MBE07	15	11	26	11	11	14	B2b	1 (0.022)
MBE08	15	12	20	10	11	15	E3a	2 (0.043)
MBE09	15	12	24	11	11	13	B2b	1 (0.022)
MBE10	15	12	24	11	11	14	B2b	1 (0.022)
MBE11[c]	15	13	21	10	11	15	BR(xB2b,DE,P)	1 (0.022)
MBE12[c]	15	13	21	10	11	15	E3a	1 (0.022)
MBE13	15	13	21	11	11	13	E3a	2 (0.043)
MBE14	15	13	24	10	11	13	B2b	1 (0.022)
MBE15[c]	15	13	25	10	11	13	B2b	1 (0.022)
MBE16[c]	15	13	25	10	11	13	E3a	1 (0.022)
MBE17	15	14	23	8	11	13	B2b	1 (0.022)
MBE18	16	12	20	10	11	15	E3a	7 (0.152)
MBE19	16	12	21	10	11	13	E3a	1 (0.022)
MBE20	16	13	21	10	11	13	E3a	1 (0.022)
MBE21	16	13	21	10	11	14	E3a	1 (0.022)
MBE22	16	13	21	10	11	15	E3a	1 (0.022)
MBE23	16	13	22	10	11	15	E3a	3 (0.065)
MBE24	16	13	24	10	11	13	B2b	1 (0.022)
MBE25	16	14	21	10	11	15	E3a	4 (0.087)
MBE26	16	14	24	10	11	13	BR(xB2b,DE,P)	1 (0.022)
MBE27	17	11	24	9	11	12	B2b	1 (0.022)
MBE28	17	12	23	11	11	14	B2b	1 (0.022)
MBE29	17	13	20	10	11	15	E3a	1 (0.022)
MBE30	17	13	21	10	11	15	E3a	1 (0.022)
MBE31	17	13	21	10	11	16	E3a	1 (0.022)
MBE32	17	14	21	10	11	15	E3a	1 (0.022)

[a]Allelic nomenclature is based on the number of repeat units.
[b]See Table 1 for definition of allelic status; nomenclature according to the Y-Chromosome consortium (2002).
[c]Haplotypes replicated across haplogroups.

Microsatellite haplotypes. Thirty different microsatellite haplotypes were detected (Table 3), with frequencies ranging from 0.022 to 0.152. The haplotypic diversity (0.966 ± 0.017) and the mean number of pairwise differences (3.246 ± 1.704) estimated are very similar to those reported for other African Pygmies (Pritchard et al., 1999; Kayser et al., 2001) and fall into the range of values reported for sub-Saharans (Kayser et al., 2001; Pritchard et al., 1999; Trovoada et al., 2001; Caglià et al., 2003; Corte-Real et al., 2000). Seven microsatellite haplotypes (23.3% of the total; haplotypes MBE05, MBE02, MBE17, MBE08, MBE10, MBE07, and MBE27 of Table 3) identified in the course of this investigation have not been observed in other sub-Saharan populations (Bosch et al., 2001; Kayser et al., 2001; Pritchard et al., 1999; Trovoada et al., 2001; Caglià et al., 2003; Corte-Real et al., 2000; a total of 772 individuals).

A percentage of 23.9 of Mbenzele haplotypes was shared by 35% of Biaka haplotypes (Pritchard et al., 1999), which represents the highest level of haplotype sharing for the Mbenzele. Five out of seven haplotypes shared between the Mbenzele and Biaka Pygmies belong to haplogroup E3a (among the Mbenzele), which suggests that they are of Bantu origin. At present, in mixed marriages between Bantu males and Pygmy females the latter usually move to the village of the husband. However, the presence of Bantu male genes in Pygmies may be explained in several ways (Cavalli-Sforza et al., 1986b): 1) admixture occurred at the beginning of the contact between farmers and Pygmies (2–3000 years ago according to Cavalli-Sforza, 1986b), facilitated by the lack of social and economic inequalities between the two groups and not necessarily constrained by a patrilocal habit; 2) children born from extramarital relationships; 3) adoption by the Pygmy communities

of orphans born from mixed marriages; or 4) return to the Pygmy society of Pygmy women and of their children after the divorce from Bantu males.

Considering non-Pygmies, the highest level of haplotype sharing occurs between the Mbenzele and Nguni speakers from Southern Africa, where 19.6% of Mbenzele haplotypes are shared with the 37.9% of haplotypes observed among the Nguni (Pritchard et al., 1999; see also Caglià et al., 2003). It is also of interest that seven out of eight Mbenzele haplotypes shared by them and Sotho Bantus from Southern Africa (Pritchard et al., 1999; see also Caglià et al., 2003) belong to haplogroup E3a. This indicates again a probable Bantu origin of these haplotypes. However, recent gene flow from the Sotho to the Mbenzele Pygmies is unlikely given the geographical distance and the lack of historical links between these two populations.

Compound haplotypes. Thirty-two compound haplotypes were detected, with frequencies ranging from 0.022 to 0.152 (Table 3). Two microsatellite haplotypes were duplicated across haplogroups, which produces a level of homoplasy of 0.067. This figure is slightly higher than that observed among Bamileke (one duplicated microsatellite haplotypes out of 19; 0.053), while no case of homoplasy was observed among the Ewondo (see Appendix A). Thomas et al. (2000) observed two duplicated haplotypes out of 142 (0.014) in a study of six microsatellites and six biallelic loci conducted among some south-African populations. Bosch et al. (1999) observed only one duplicated haplotype out of 56 (0.017) in a survey conducted in northwestern Africa by means of 11 microsatellites and six biallelic loci.

The allele frequencies at the six microsatellite loci examined within haplogroups with more than one observation are shown in Table 4. The allele frequencies within the two most frequent haplogroups (E3a and B2b), vary markedly, with DYS392 being the only exception. A reliable comparison between the Mbenzele and the Bamileke and Ewondo is possible only for haplogroup E3a (see Appendix B). The three populations show the same prevalent allele at four out of six loci:

TABLE 4. Single-locus frequency of microsatellite alleles for each haplogroup in the Mbenzele Pygmies[a]

		HG B2b (n = 11)	HG E3a (n = 28)	HG E(xE3a) (n = 4)	Overall (46)
	Allele				
DYS19	14	—	—	1	0.109
	15	0.727	0.214	—	0.326
	16	0.091	0.643	—	0.435
	17	0.182	0.143	—	0.130
DYS389I	11	0.364	—	—	0.087
	12	0.273	0.357	0.250	0.304
	13	0.273	0.464	0.750	0.457
	14	0.091	0.179	—	0.152
DYS390	20	—	0.357	—	0.217
	21	—	0.500	—	0.348
	22	—	0.107	—	0.065
	23	0.182	—	—	0.043
	24	0.545	—	1	0.239
	25	0.182	0.036	—	0.065
	26	0.091	—	—	0.022
DYS391	8	0.091	—	—	0.022
	9	0.091	—	—	0.022
	10	0.273	0.929	0.250	0.717
	11	0.545	0.071	0.750	0.239
DYS392	11	1	1	1	1.000
DYS393	12	0.091	—	—	0.022
	13	0.455	0.179	1	0.326
	14	0.455	0.036	—	0.152
	15	—	0.750	—	0.478
	16	—	0.036	—	0.022

[a]The haplogroups P and BR(xB2b,DE,P) were not considered separately since they were found in only one or two individuals, respectively.

DYS389I (allele13), DYS390 (allele 21), DYS391 (allele 10), and DYS392 (allele 11). The relative similarity among allelic frequencies within haplogroups is also reflected by the AMOVA, which showed that the percentage fraction of total variation attributable to differences among haplogroups (56.6%, $P < 0.001$) greatly exceeds that attributable to differences among populations (8.9%, $P < 0.001$). Therefore, our results extend to sub-Saharan populations the notion that Y-chromosomal microsatellite variation is deeply structured by haplogroups (Bosch et al., 1999).

Genetic relationships among the Mbenzele and other sub-Saharan populations

Comparison of SNPs and microsatellites. It has been shown that Y-chromosome haplotypes defined by microsatellites that are shared by individuals and/or populations can be identical only by state but not by descent (De Knijff, 2000). This may confound population relationships so that the interpopulational diversity inferred from microsatellite loci might not reflect the true genetic and historical relationships among populations (see Parra et al., 1999). However, as shown by Bosch et al. (1999) and confirmed by our study, the microsatellite variation is deeply structured by haplogroups so that only a minor part of microsatellite haplotypes is duplicated across haplogroups. Thus, microsatellite variation should reflect, at least in part, that of haplogroups and retain some phylogenetic information.

In a previous study of Y-chromosome microsatellites we compared the capability of F_{st} and R_{st} genetic distances to produce results congruent with a priori expectations based on anthropobiological and/or ethnohistorical knowledge of some sub-Saharan populations (Caglià et al., 2003). It was observed that the F_{st} provides a more reliable picture of differentiation among sub-Saharan populations than the R_{st} genetic distance. In this study we compare patterns of interpopulational diversity based on microsatellite data with that based on haplogroup frequencies. A total of six populations were available for this analysis, the Mbenzele Pygmies and Biaka Pygmies, the Bamileke and Ewondo from Cameroon, the Eastern Pygmies from Zaire, and the !Kung from Southern Africa (see the legend of Fig. 1 for data sources).

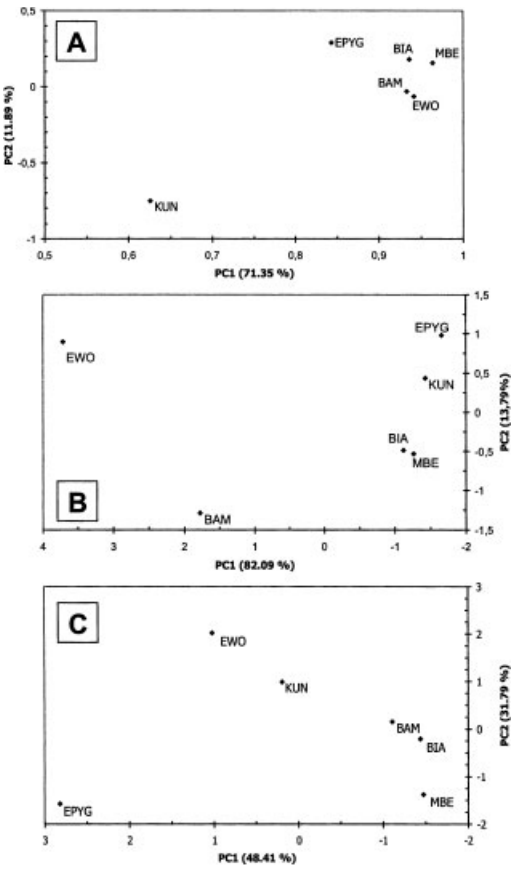

Fig. 1. Principal component plots of haplogroup frequencies (**A**), F_{st} genetic distances based on microsatellite data (**B**) and R_{st} genetic distances based on microsatellite data (**C**). List of abbreviations and data source for haplogroup frequencies: BAM, Bamileke (Cameroon; this study); BIA, Biaka Pygmies (Central African Republic; Cruciani et al., 2002); EWO, Ewondo (Cameroon; this study); KUN, !Kung (South Africa; Cruciani et al., 2002); MBE, Mbenzele Pygmies (Central African Republic; this study); EPYG, Eastern Pygmies (Democratic Republic of Congo, Cruciani et al., 2002). List of abbreviations and data source for microsatellite data: BAM, Bamileke (Cameroon; Caglià et al., 2003); BIA, Biaka Pygmies (Central African Republic; Pritchard et al., 1999; http://pritch.bsd.uchicago.edu/; 20 chromosomes); EWO, Ewondo (Cameroon; Caglià et al., 2003); KUN, !Kung (South Africa; Pritchard et al., 1999; http://pritch.bsd.uchicago.edu/; 29 chromosomes); MBE, Mbenzele Pygmies (Central African Republic; this study); EPYG, Eastern Mbuti Pygmies, Zaire (Democratic Republic of Congo, http://pritch.bsd.uchicago.edu/;). The !Kung sample was obtained by pooling 14 Sekele San and 15 Omega San, whose difference in haplotype distribution was found to be statistically insignificant using the exact test implemented in the Arlequin software (Schneider et al., 1997).

The PC plots are shown in Figure 1. The high level of variance accounted for by the two

first components (from 95.9% for that based on F_{st} genetic distances to 80.2% for that based on R_{st} genetic distances) indicates that the plots provide a satisfactory representation of the original data. In the plot based on haplogroup frequencies the Mbenzele are close to the Biaka Pygmies, while the Bamileke cluster even more tightly with the Ewondo. These results reflect the common affiliation and geographical position of the two Western Pygmy populations and confirm the similarity between the two Bantu groups observed in previous studies of protein loci and microsatellite variation (Spedini et al., 1999; Destro-Bisol et al., 2000). The Eastern Pygmies do not cluster with the two Western Pygmy populations (Biaka and Mbenzele). The !Kung are an outlier, a result mainly due to the high frequency of haplogroup A, which does not occur in any other population. The divergence of the !Kung from other sub-Saharans noted here is consistent with the uniqueness of some of their external characters in the sub-Saharan context (e.g., yellowish skin, steatopygy, and epichantic eyefolds) and with genetic distances based on data of mtDNA (Vigilant et al., 1991, Chen et al., 1995, 2000) and protein loci (Cavalli-Sforza et al., 1994). In the F_{st}-based plot the clustering between the Biaka and Mbenzele is replicated and the association between Eastern and Western Pygmies is even weaker than in the plot based on haplogroup frequencies. In fact, the Biaka and Mbenzele are more similar to the !Kung than to the Eastern Pygmies. The Bamileke and Ewondo are less similar than in the PC plot of haplogroup frequencies, but a certain separation from the remaining populations is maintained. It is of interest that the first component which accounts for 82.1% of total variance separates very effectively hunter-gatherers, Pygmies, and !Kung from Bantu-speaking populations. The Pygmies and !Kung are thought to be in genetic continuity with the ancient inhabitants of Africa to the south of the equatorial belt, whereas the settlement of Bantu-speaking populations in this area is estimated to have occurred not earlier than 3–4000 years ago (Cavalli-Sforza et al., 1994; Vigilant et al., 1989; Vansina, 1984). The spatial relationships among populations in the plot based on R_{st} genetic distances does not give easily understandable information regarding their history. The only exception is in the clustering between the Biaka and Mbenzele, which was, however, been observed only in the first principal component, which accounts for only 48.4% of total variance.

To sum up, our comparison highlights some interesting differences between the plot built using haplogroup frequencies and those based on genetic distances calculated from microsatellite haplotypes. The F_{st}-based plot provides a picture of population relationships which is in part congruent and in part complementary to that based on haplogroup frequencies. This result strengthens our previous conclusion (Caglià et al., 2003) that Y-chromosome microsatellite data may provide useful information for analyses of genetic relationships among sub-Saharan populations.

Comparison of Y-chromosomal and autosomal microsatellites. Y-chromosomal markers are subjected to a higher pressure by genetic drift than autosomal microsatellites owing to the haploid mode of inheritance, which reduces their effective population size to one-fourth of that of autosomes. This important difference should also have an effect on the genetic distances among populations, with a tendency of Y-chromosomal microsatellites to show larger interpopulational differentiation than autosomal microsatellites when genetic drift is the main agent differentiating populations. A comparison between the two genetic systems is possible in this study because five out of six populations used for the comparison of microsatellites and SNPs of Y-chromosome have been studied for six autosomal microsatellites (Destro-Bisol et al., 2000; Pérez-Lezaun et al., 1997). Although we cannot exclude that part of the difference is due to random processes, since not all Y-chromosomal and autosomal results were obtained using the same individuals, and as the number of loci was limited, there is a substantial difference between the two genetic systems which merits attention. In fact, the Mbenzele and Biaka Pygmies are separated by a larger genetic distance in the PC plot built using autosomal microsatellites (Fig. 2C). This finding is in contrast with the expectation of the greater effect of genetic drift on Y-chromosome markers but is in agreement with the common affiliation of the Mbenzele and Biaka to the western cluster of African Pygmies (Cavalli-Sforza, 1986a). It should be noted that the plot based on Y-chromosome frequencies of single loci (Fig. 2B) reconfirms

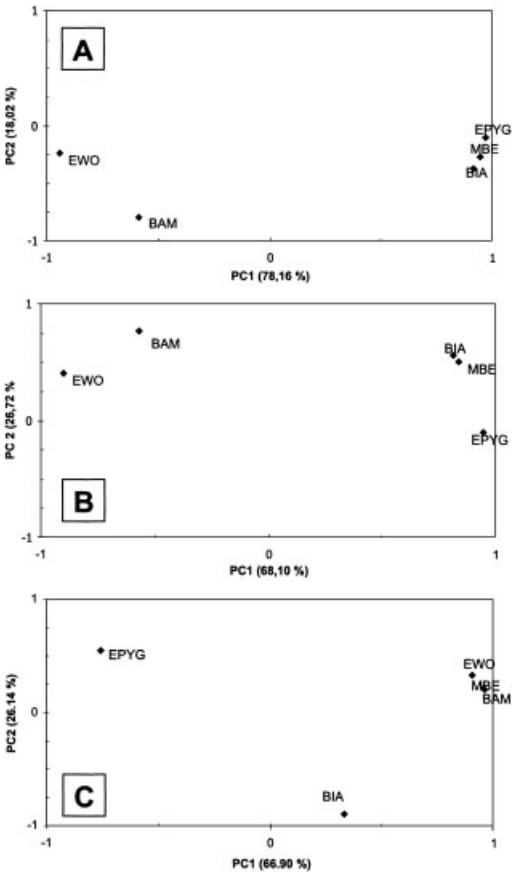

Fig. 2. Principal component plots F_{st} genetic distances based on frequencies of Y-chromosome haplotypes (**A**), allele frequencies at loci DYS19, DYS389I, DYS390, DYS391, DYS392, and DYS393 (**B**), and alleles at autosomal loci D3S1358, F13A1, FES, TH01, and VWA (**C**). Data source for allele frequencies at autosomal loci: Bamileke (92 chromosomes: Destro-Bisol et al., 2000); Biaka Pygmies (20 chromosomes: Perez-Lezaun et al., 1997); Ewondo (130 chromosomes: Destro-Bisol et al., 2000); Mbenzele Pygmies (96–98 chromosomes: Destro-Bisol et al., 2000); Eastern Mbuti Pygmies (20 chromosomes: Perez-Lezaun et al., 1997).

the close genetic similarity between the two Western Pygmy populations. This indicates that the discrepancy between the two genetic systems is not simply due to the computational consequences of different levels of within- and between-population diversity estimated by using allelic and haplotypic frequencies. We believe that two, mutually compatible, biological explanations should be taken into consideration to explain the different behavior of Y-chromosomal and autosomal microsatellites. First, the lack of recombination in most of the Y-chromosome (including that where the microsatellite loci used in this study are located) makes possible the accumulation of mutation along lines of descent. This property allows Y-chromosomes to maintain more phylogenetic information than autosomes. This also influences the variation of microsatellite loci (Bosch et al., 1999) and probably contributes to the better agreement of the Y-chromosome microsatellite variation with expected population relationships (i.e., the close similarity between the two Western Pygmy populations), at least when the comparison is made using a similar and limited number of loci. Second, a certain contribution to the relative similarity between the Mbenzele and Biaka could come from the fact that both of these populations have been influenced by the male-driven gene flow during the Bantu expansion (Hammer et al., 2001; Cruciani et al., 2002). Unfortunately, the Bamileke and Ewondo are of little use to test this hypothesis since they are of Sudanic origin and acquired the Bantu language much later than the Bantu expansion (Spedini et al., 1999). However, as described above, it is significant that five out of seven haplotypes shared between the Mbenzele and Biaka Pygmies are associated with a Bantu marker such as the M2 mutation defining haplogroup E3a.

ACKNOWLEDGMENTS

We thank the *Equipe du Dispensaire de Santé Integrée, Mission Catholique du Belemboké* (Central African Republic) and the blood donors, whose availability made this study possible. We are indebted to Cristian Capelli (Catholic University of Rome) for constructive comments. This research was supported by grants from the M.U.R.S.T. (Cofin Project "Dna e Biodemografia: approccio integrato allo studio della mobilità umana", protocol 2003054059-002 and 2003054059-005) to G.P. and G.D.B. and the University of Rome "La Sapienza" (Ricerche di Ateneo) to G.D.B.

LITERATURE CITED

Bosch E, Calafell F, Santos FR, Perez-Lezaun A, Comas D, Benchemsi N, Tyler-Smith C, Bertranpetit J. 1999. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. Am J Hum Genet 65:1623–1638.

Bosch E, Calafell F, Comas D, Oefner PJ, Underhill PA, Bertranpetit J. 2001. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian Peninsula. Am J Hum Genet 68: 1019–1029.

Caglià A, Tofanelli S, Coia V, Boschi I, Pescarmona I, Spedini G, Pascali V, Paoli G, Destro-Bisol G. 2003. A study of Y-chromosome microsatellite variation in sub-Saharan Africa: a comparison between F_{st} and R_{st} genetic distances. Hum Biol 75:310–330.

Cavalli-Sforza LL. 1986a. Demographic data. In: Cavalli-Sforza LL, editor. African Pygmies. Orlando, FL: Academic Press. p 23–44.

Cavalli-Sforza LL. 1986b. African Pygmies: an evaluation of the state of research. In: Cavalli-Sforza LL, editor. African Pygmies. Orlando, FL: Academic Press. p 361–426.

Cavalli-Sforza L, Menozzi P, Piazza A. 1994. The history and geography of human genes. Princeton, NJ: Princeton University Press.

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.

Chen YS, Olckers A, Schurr TG, Kogelnik AM, Huoponen K, Wallace DC. 2000. mtDNA variation in the South African Kung and Khwe—and their genetic relationships to other African populations. Am J Hum Genet 66:1362–1383.

Coia V, Sansonetti B, Paoli G, Tofanelli S, Spedini G, Destro-Bisol G. 2002. Variation at 10 protein coding loci in the Mbenzele Pygmies from the Central African Republic and a comparison with microsatellite data. Am J Hum Biol 14:9–14.

Corte-Real F, Carvalho L, Andrade L, Anjos MJ, Pestoni C, Lareu MV, Carracedo A, Vieira DN, Vide MC. 2000. Chromosome Y STRs analysis and evolutionary aspects for Portuguese speaking countries. In: Sensabaugh GF, Lincoln PJ, Olaisen B, editors. Progress in forensic genetics 8. Amsterdam: Elsevier. p 272–277.

Corvin H. 1988. Les Pygmees Mbenzele en quete de la maitrise de leur devenir: Analyse d'une mutation social a partir de l'autoévaluation d'un projet de développement. Diplòme des Hautes Etudes de la Pratique Sociale, Centre de formation pour le développement. Université de Strasbourg II.

Cruciani F, Santolamazza P, Shen P, Macaulay V, Moral P, Olckers A, Modiano D, Holmes S, Destro-Bisol G, Coia V, Wallace DC, Oefner PJ, Torroni A, Cavalli-Sforza LL, Scozzari R, Underhill PA. 2002. An Asia to Sub-Saharan Africa back migration is supported by high-resolution analysis of human Y chromosomes. Am J Hum Genet 70:1197–1214.

De Knijff P. 2000. Y chromosomes shared by descent or by state. In: Renfrew C, Boyle K, editors. Archaeogenetics: DNA and the population prehistory of Europe. Cambridge: McDonald Institute. p 301–304.

Destro-Bisol G, Belledi M, Capelli C, Maviglia R, Spedini G. 2000a. Genetic variation at the ApoB 3′ HVR minisatellite locus in the Mbenzele Pygmies from the Central African Republic. Am J Hum Biol 12:588–592.

Destro-Bisol G, Boschi I, Caglià A, Tofanelli S, Pascali V, Paoli G, Spedini G. 2000b. Microsatellite variation in Central Africa: an analysis of intrapopulational and interpopulational genetic diversity. Am J Phys Anthropol 112:319–337.

Foster EA, Jobling MA, Taylor PG, Donnelly P, de Knijff P, Mieremet R, Zerjal T, Tyler-Smith C. 1998. Jefferson fathered slave's last child. Nature 396:27–28.

Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, Zegura SL. 1997. The geographic distribution of human Y chromosome variation. Genetics 145:787–805.

Hammer MF, Karafet T, Rasanayagam A, Wood ET, Altheide TK, Jenkins T, Griffiths RC, Templeton AR, Zegura SL. 1998. Out of Africa and back again: nested cladistic analysis of human Y chromosome variation. Mol Biol Evol 15:427–441.

Hammer MF, Karafet TA, Redd AJ, Jarjanazi H, Santachiara-Benerecetti S, Soodyall H, Zegura SL. 2001. Hierarchical patterns of global human Y-chromosome diversity. Mol Biol Evol 18:1189–1203.

Jobling MA, Pandya A, Tyler Smith C. 1997. The Y chromosome in forensic analysis and paternity testing. Int J Legal Med 110:118–124.

Kayser M, Krawczak M, Excoffier L, Dieltjes P, Corach D, Pascali V, Gehrig C, Bernini LF, Jespersen J, Bakker E, Roewer L, de Knijff P. 2001. An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. Am J Hum Genet 68: 990–1018.

Michalakis Y, Excoffier L. 1996. A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. Genetics 142:1061–1064.

Murdock GP. 1959. Africa: its people and their culture history. New York: McGraw-Hill.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Paracchini S, Arredi B, Chalk R, Tyler-Smith C. 2002. Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. Nucleic Acids Res 30:e27.

Parra E, Shriver MD, Soemantri A, McGarvey ST, Hundrieser J, Saha N, Deka R. 1999. Analysis of five Y-specific microsatellite loci in Asian and Pacific populations. Am J Phys Anthropol 110:1–16.

Perez-Lezaun A, Calafell F, Mateu E, Comas D, Bosch E, Bertranpetit J. 1997. Allele frequencies for 20 microsatellites in a worldwide population Survey. Hum Hered 47:189–196.

Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. Mol Biol Evol 16:1791–1798.

Reynolds J, Weir BS, Cockerham CC. 1983. Estimation of the coancestry coefficient: basis for a short term genetic distance. Genetics 105:767–779.

Roewer L, Krawczak M, Willuweit S, Nagy M, Alves C, Amorim A, Anslinger K, Augustin C, Betz A, Bosch E, Caglià A, Carracedo A, Corach D, Dekairelle AF, Dobosz T, Dupuy BM, Furedi S, Gehrig C, Gusmao L, Henke J, Henke L, Hidding M, Hohoff C, Hoste B, Jobling MA, Kargel HJ, de Knijff P, Lessig R, Liebeherr E, Lorente M, Martinez-Jarreta B, Nievas P, Nowak M, Parson W, Pascali VL, Penacino G, Ploski R, Rolf B, Sala A, Schmidt U, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Kayser M. 2001. Online reference database of European Y-chromosomal short tandem repeat (STR) haplotypes. Forensic Sci Int 118: 106–113.

Schneider S, Kueffer JM, Roessli D, Excoffier L. 1997. Arlequin ver1.1: a software for population genetic data analysis. Genetics and Biometry Laboratory. University of Geneva, Switzerland.

Scozzari R, Cruciani F, Santolamazza P, Malaspina P, Torroni A, Sellitto D, Arredi B, Destro-Bisol G, De Stefano G, Rickards O, Martinez-Labarga C, Modiano D, Biondi G, Moral P, Olckers A, Wallace DC, Novelletto A. 1999. Combined use of biallelic and micro-

satellite Y-chromosome polymorphisms to infer affinities among African populations. Am J Hum Genet 65:829–846.
Slatkin M. 1995. A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462.
Spedini G, Destro-Bisol G, Mondovi S, Kaptue L, Taglioli L, Paoli G. 1999. The peopling of sub-Saharan africa: the case study of Cameroon. Am J Phys Anthropol 110:143–162.
Thomas MG, Skorecki K, Ben-Ami H, Parfitt T, Bradman N, Goldstain DB. 1998. Origins of Old Testament priests. Nature 394:138–140.
Thomas MG, Parfitt T, Weiss DA, Skorecki K, Wilson JF, le Roux M, Bradman N, Goldstein DB. 2000. Y chromosomes traveling south: the Cohen modal haplotype and the origins of the Lemba, the "Black Jews of Southern Africa." Am J Hum Genet 66: 674–686.
Trovoada MJ, Alves C, Gusmao L, Abade A, Amorim A, Prata MJ. 2001. Evidence for population sub-structuring in Sao Tome e Principe as inferred from Y-chromosome STR analysis. Ann Hum Genet 65:271–283.
Underhill PA, Passarino G, Lin AA, Shen P, Mirazon Lahr M, Foley RA, Oefner PJ, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. Ann Hum Genet 65:43–62.
Vansina J 1984. Western Bantu expansion. J Afr Hist 25:129–145.
Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. 1989. Mitochondrial DNA sequences in single hairs from a southern African population. Proc Natl Acad Sci USA 86:9350–9354.
Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. 1991. African populations and the evolution of human mitochondrial DNA. Science 253:1503–1507.
Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370.
Y Chromosome Consortium (YCC). 2002. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome Res 12:339–348.

APPENDIX

A. Compound Y-chromosome haplotypes of the Bamileke (49 individuals) and Ewondo (25 individuals) from Cameroon

Compound haplotype	Microsatellite haplotype[a]						HG[b]	N (% frequency)	
	DYS19	DYS389I	DYS390	DYS391	DYS392	DYS393		Bamileke	Ewondo
BAM01[c]	14	12	25	10	11	13	E(xE3a)	1 (0.020)	—
BAM02[c]	14	12	25	11	11	13	E3a	1 (0.020)	—
BAM03	14	12	25	11	11	13	E(xE3a)	2 (0.041)	—
BAM04	15	11	21	10	11	13	E3a	16 (0.327)	—
BAM05	15	12	21	10	11	13	E3a	1 (0.020)	—
EWO01	15	13	21	10	11	13	E3a	6 (0.122)	7 (0.280)
EWO02	15	13	21	10	11	14	E3a	—	1 (0.040)
EWO03	15	13	21	10	11	15	E3a	1 (0.020)	1 (0.040)
EWO04	15	13	24	11	11	14	BR(xB2b,DE,P)	—	1 (0.040)
EWO05	15	14	21	10	11	13	E3a	—	8 (0.320)
EWO06	15	14	21	10	11	14	E3a	—	1 (0.040)
EWO07	15	14	21	10	11	15	E3a	—	1 (0.040)
BAM06	15	14	21	11	11	13	E3a	1 (0.020)	—
BAM07	16	11	21	10	11	13	E3a	1 (0.020)	—
EWO08	16	13	21	10	11	13	E3a	1 (0.020)	1 (0.040)
BAM08	16	13	21	10	11	14	E3a	2 (0.041)	—
BAM09	16	13	21	10	11	15	E3a	3 (0.061)	1 (0.040)
BAM10	16	13	21	10	11	16	E3a	1 (0.020)	—
BAM11	16	13	21	11	11	13	E3a	1 (0.020)	—
BAM12	16	13	22	11	11	15	E3a	1 (0.020)	—
BAM13	16	14	21	10	11	14	E3a	1 (0.020)	—
BAM14	16	14	24	10	13	13	P	3 (0.061)	—
BAM15	17	13	21	10	11	14	E3a	3 (0.061)	—
EWO10	17	13	21	10	11	15	E3a	—	1 (0.040)
BAM16	17	14	21	10	11	14	E3a	1 (0.020)	—
BAM17	17	14	21	10	11	15	E3a	2 (0.041)	—
EWO11	17	14	24	10	11	13	BR(xB2b,DE,P)	—	2 (0.080)

[a]Allelic nomenclature is based on the number of repeat units.
[b]See Table 1 for definition of allelic status nomenclature according to the Y-chromosome consortium (2002).
[c]Haplotypes replicated across haplogroups.

B. Single-locus frequency of microsatellite alleles for each haplogroup in the Bamileke and Ewondo

		HG P	HG BR(xB2b,DE,P)	HG E3a		HG E(xE3a)
		Bamileke (n = 3)	Ewondo (n = 3)	Bamileke (n = 43)	Ewondo (n = 22)	Bamileke (n = 3)
Locus	Allele					
DYS19	14	—	—	0.023	—	1
	15	—	0.333	0.581	0.864	—
	16	1	—	0.256	0.091	—
	17	—	0.667	0.140	0.045	—
DYS389I	11	—	—	0.395	—	—
	12	—	—	0.047	—	1
	13	—	0.333	0.442	0.545	—
	14	1	0.667	0.116	0.455	—
DYS390	21	—	—	0.953	1	—
	22	—	—	0.023	—	—
	24	1	1	—	—	—
	25	—	—	0.023	—	1
DYS391	10	1	0.667	0.907	1	0.333
	11	—	0.333	0.093	—	0.667
DYS392	11	—	1	1	1	1
	13	1	—	—	—	—
DYS393	13	1	0.667	0.651	0.727	1
	14	—	0.333	0.163	0.091	—
	15	—	—	0.163	0.182	—
	16	—	—	0.023	—	—