Korean J Biol Sci 2: 281-285, 1998

PCR-Based Polymorphic Analysis for the Y Chromosomal Loci *DYS19* and *DXYS5Y* (47z) in the Korean Population

Dong Jik Shin, Yung Jin Kim[1], and Wook Kim*

Department of Biology, Dankook University, Cheonan 330-714, Korea;
[1]Department of Biology, Chungnam National University, Taejon 305-764, Korea

Key Words:
Human Y chromosome
DYS19
DXYS5Y
Japanese population
Korean population

We examined Y chromosomal DNA polymorphisms at the *DYS19* and *DXYS5Y* loci in a total of 480 unrelated male samples from the Korean population. All five common alleles were identified at the tetranucleotide microsatellite locus *DYS19* in this study. The C allele was the most frequent (212/480), followed by D (136/480), B (75/480), E (36/480) and A (21/480) allele. The frequency of Y2 allele at the *DXYS5Y* locus was found to be 4.6% (22/480). Combining the allelic variation at these two loci resulted in a total of 9 combination haplotypes. The mean combination haplotype diversity was 0.72. Based on the results of these two loci, Korean and Japanese populations may share some common genetic structure that is rare or absent in the other ethnic groups. The genetic similarity between Korean and Japanese populations may be due to the large infusion of Y chromosomes through the Yayoi migration starting 2,300 years ago from Korea to Japan.

The human Y chromosome is inherited in a patrilineal manner and does not undergo recombination with the exception of the pseudoautosomal region. Therefore, the Y chromosome retains a unique record of mutational events that occurred in previous generations, and Y haplotypes constructed with multiple polymorphic sites can be used for studies of paternal lineages and population history in man, just as the mitochondrial DNA has come to be for female lineages (Ngo et al., 1986; Cann et al., 1987; Hammer, 1994; Spurdle et al., 1994). The polymerase chain reaction (PCR) has also facilitated detection of DNA sequence variation on the Y chromosome (Edwards et al., 1991; Hammer, 1995; Jobling and Tyler-Smith, 1995).

The human Y-linked short tandem repeats (STRs) or microsatellite, *DYS19*, containing a GATA motif has been widely used as a marker for studies of the structure and dynamics of human populations (Roewer et al., 1992; Santos, 1993, 1996; Gomolka et al., 1994). There are at least nine different alleles including five common alleles (A-E) which exhibit remarkable and frequent differences in worldwide populations (Santos et al., 1993, 1996; Gomolka et al., 1994; Hammer and Horai, 1995; Hammer et al., 1997). Based on Southern analysis, the probe p47z detects variations at loci on the short arm of Y (*DXYS5Y*) and the long arm of the X (*DXYS5X*) chromosomes (Nakahori et al., 1989; Nakagome et al., 1992). Allelic variation of the *DXYS5Y* now can be scored by PCR, and increased speed of the PCR offers advantages over Southern analysis for studies of the human population and personal identification. PCR amplification using specific primers and the subsequent digestion of PCR products with the restriction endonuclease *Stu*I identify a two-allele, Y1 and Y2 in the locus *DXYS5Y* (M. Hammer, personal communication). Here, we report the frequencies of these two Y-linked specific markers in the Korean population to compare the genetic structure among different human ethnic groups.

* To whom correspondence should be addressed.
Tel: 82-417-550-3441, Fax: 82-417-550-3441

Materials and Methods

Subjects and DNA extraction

Buccal cells were collected from a total of 480 unrelated Korean males for the detection of allelic variations at the loci *DYS19* and *DXYS5Y*. DNA was isolated from buccal cells by following the procedure of Richards et al. (1993).

PCR amplification and allele designation

The *DYS19* alleles were amplified by PCR using flanking primers as described by Hammer and Horai (1995). The DNA fragments derived from PCR amplification of the *DYS19* locus were separated by polyacrylamide gel electrophoresis (PAGE) for 10 h at 100 V. Following the PAGE, the gel (10%) was stained with ethidium bromide and the fragments were visualized

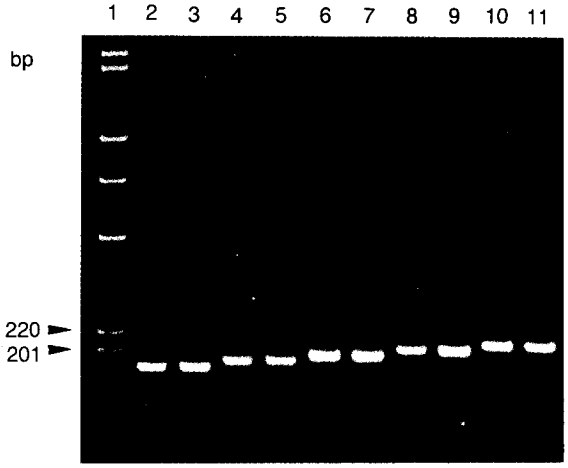

Fig. 1. Polyacrylamide gel electrophoresis of PCR products amplified with specific flanking primers of the *DYS19*. Ten individuals displaying different allelic classes at the *DYS19* locus in increasing order of size. 1, DNA size markers. 2, 3, 186 bp. 4, 5, 190 bp. 6, 7, 194 bp. 8, 9, 198 bp. 10, 11, 202 bp.

by UV light. We found five common alleles (A-E) ranging in size from 186 bp to 202 bp in four-nucleotide (4-nt) increments (Fig. 1).

The *DXYS5Y* locus was amplified by PCR using primers designed to flank the *DXYS5Y* (M. Hammer,

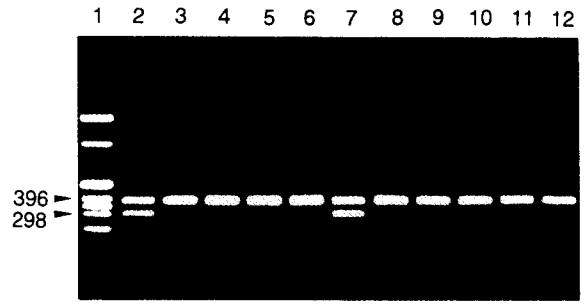

Fig. 2. Detection of the allelic variation for the locus *DXYS5Y* in PCR products. Electrophoresis in a 2% agarose gel of restriction fragments of the PCR product produced by digestion with *StuI*. 1, DNA size markers. 2, 7, Y2 allele samples. 3~6, 8~12, Y1 allele samples.

personal communication): Primer 1, 5′-TAGTTACGCC-TTGGCATAAC-3′; Primer 2, 5′-TGAGTCAATGTCAAT-GAATC-3′. The reactions were performed in a 25 μl final volume containing 50 ng genomic DNA, 5 pM primers, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 50 mM KCl, Tris-HCl (pH 8.3) and 1.25 U Ampli*Taq* DNA polymerase (Perkin-Elmer). The amplification was performed at 94℃ for 3 min, and then 35 cycles at 94℃ for 30 sec, 53℃ for 45 sec, and 72℃ for 45 sec. The PCR products were digested with *StuI* and resolved on 2% agarose gel. There were two copies of *DXYS5* amplified by

Table 1. Frequencies of the *DYS19* alleles in several Asian populations

Population (n)	No. (%) of							
	Z	A	B	C	D	E	F	x/y
North Asians[a, b]								
Siberians (503)*	7 (1.4)	104 (20.7)	207 (41.2)	81 (16.1)	82 (16.3)	21 (4.2)	0	1 (0.2)
North Chinese (64)	0	5 (7.8)	3 (4.7)	14 (21.9)	29 (45.3)	13 (20.3)	0	0
Central Asians[a, b]								
Altai (29)	0	3 (10.3)	1 (3.4)	4 (13.8)	20 (69.0)	0	0	1 (3.4)
Mongolians (148)	0	3 (2.0)	33 (22.3)	46 (31.1)	41 (27.7)	13 (8.8)	1 (0.7)	11 (7.4)
Tibetans (34)	0	0	7 (20.6)	23 (67.6)	4 (11.8)	0	0	0
East Asians								
Japanese (132)[a]	0	10 (7.6)	4 (3.0)	65 (49.2)	33 (25.0)	20 (15.2)	0	0
Koreans (480)[c]	0	21 (4.4)	75 (15.6)	212 (44.2)	136 (28.3)	36 (7.5)	0	0
Chinese (121)**, [b]	0	4 (3.3)	22 (18.2)	48 (39.7)	27 (22.3)	19 (15.7)	0	1 (0.8)
Taiwanese (77)***, [b]	1 (1.3)	4 (5.2)	20 (26.0)	30 (39.0)	17 (22.1)	5 (6.5)	0	0
Southeast Asians[a, b]								
Indonesians (97)	0	2 (2.1)	22 (22.7)	46 (47.4)	18 (18.6)	9 (9.3)	0	0
South Asians[a, b]								
Indians (43)	0	1 (2.3)	15 (34.9)	17 (39.5)	8 (18.6)	2 (4.7)	0	0
Sri Lankans (87)	0	2 (2.3)	25 (28.7)	41 (47.1)	14 (16.1)	5 (5.7)	0	0
Pakistanis (25)	0	3 (12.0)	3 (12.0)	10 (40.0)	8 (32.0)	1 (4.0)	0	0
Oceanians[a, b]								
Aus. Aborigins (82)	0	0	15 (18.3)	32 (39.0)	24 (29.3)	11 (13.4)	0	0
Pa. N. Guineans (48)	0	0	16 (33.3)	22 (45.8)	7 (14.6)	3 (6.3)	0	0

[a] Hammer et al. (1997); [b] Hammer and Karafet (personal communication); [c] Present study.
* North central Asian, largely in Russia (Hammer and Karafet, personal communication).
** Most are southern mainland Chinese (Hammer et al., 1997).
*** Most are Chinese ancestry in Taiwan (Hammer and Hoari, 1995).

PCR: one copy on X and the other on Y chromosome. In addition, there were two bands after digestion for Y2 allele of the *DXYS5Y*, because the PCR fragment from the Y chromosome cuts only with *StuI* (Fig. 2).

Data analysis

The mean haplotype (gene) diversity was calculated for each group as $h = n\ (1 - \sum x_i^2)/(n-1)$, where n represents the number of chromosomes sampled and x_i is the frequency of the ith haplotype (Nei, 1987). The PHYLIP Package (Felsenstein, 1993) was used to compute genetic distances for use in distance matrix programs and for constructing maximum likehood phylogeny estimates based on Y chromosome haplotype frequencies. Genetic distances were calculated using Cavalli-Sforza's Chord genetic distance ($4D$) (Cavalli-Sforza and Bodmer, 1971). The neighbor-joining (NJ) (Saitou and Nei, 1987) was used to construct branching diagrams from matrices of pairwise distances.

Results

Variation at the DYS19 locus

In earlier surveys, five alleles (A-E) corresponding to the 186 bp to 202 bp fragments were found to be common in human populations (Roewer et al., 1992; Santos et al., 1993; Gomolka et al., 1994; Muller et al., 1994; Hammer et al., 1997). Recently, three rare allele classes have been reported: the 178 bp (178 bp allele) (Ciminelli et al., 1995), the 182 bp (Z allele) (Santos et al., 1996), and the 206 bp (F allele) (Hammer and Horai, 1995).

In this survey of alleles at the *DYS19* locus, all five common alleles (A-E) were identified in the Korean population (Table 1). The C allele was the most frequent (44.2%), followed by the D (28.3%), B (15.6%), E (7.5%), and A (4.4%) alleles. The predominance of the C allele in this study was similar to previous reports from other surveys of Asian, Australian, and African populations. The mean haplotype diversity of the *DYS19* was 0.694 in this study. This *DYS19* allelic diversity (0.694) appeared to be nearly the same as that of other Asian populations (0.695) (Hammer et al., 1997). Cavalli-Sforza's Chord genetic distance based on the frequencies for *DYS19* alleles (Table 1) showed that Japanese have a closer genetic relationship ($4D$= 0.0198) with Koreans than any other Asian populations.

Variation at the DXYS5Y locus

The frequency of the *DXYS5Y* alleles determined in this study was compared with the results of previous surveys (Table 2). The Y2 allele frequency in Korea was found to be 4.6% (22/480). The frequency of Y2 allele in this study was somewhat lower than in the previous results of Nakagome et al. (1992): the Y2 allele was present in 4 of 41 males (9.8%) in Korea.

Table 2. Frequencies of the *DXYS5Y* alleles in several Asian populations

Population (n)	No. (%) of	
	Y1	Y2
North Asians[a, b]		
North Chinese (86)	86 (100)	0
Central Asians[b]		
Altai (29)	29 (100)	0
Mongolians (30)	41 (100)	0
Tibetans (22)	22 (100)	0
East Asians		
Japanese (283)[a, b, c]	217 (76.5)	66 (23.3)
Koreans (480)[d]	458 (95.4)	22 (4.6)
Chinese (144)[a, b]	144 (100)	0
Taiwanese (44)[b]	44 (100)	0
Southeast Asians[b]		
Indonesians (10)	10 (100)	0
South Asians[b]		
Indians (19)	19 (100)	0
Pakistanis (25)	17 (100)	0
Oceanians[b]		
Australian Aborigins (29)	29 (100)	0

[a] Nakagome et al. (1992); [b] Hammer and Karafet (personal communication);
[c] Hammer and Horai (1995); [d] Present study.

However, the result for the presence of Y2 allele in this survey is consistent with the previous reports that the Y2 allele is restricted to Korea and Japan (Nakagome et al., 1992; Mathias et al., 1994; Hammer and Horai, 1995).

DYS19/DXYS5Y combination haplotypes and population tree

A total of nine haplotypes were found with respect to the *DYS19/DXYS5Y* combination (Table 3). The mean combination haplotype diversity was 0.72. The haplotype H3 (C/Y1) was the most frequent (42.1%), and the frequencies of combination haplotypes were proportional to the frequencies of the *DYS19* (Tables 1, 3). Each chromosome of Y1 and Y2 was associated with all of 5 *DYS19* alleles with the exception of the combination of A/Y2.

Cavalli-Sforza's Chord genetic distance-NJ tree (Fig. 3) based on the frequencies of *DYS19* (Table 1) and *DXYS5Y* (Table 2) alleles showed that Japanese and Koreans have a closer genetic relationship with Chinese and Mongolians than with other Asians.

Discussion

Based on results of the frequency of the *DYS19* alleles, Japanese show a closer genetic relationship with Koreans than with other Asians (Table 1). This

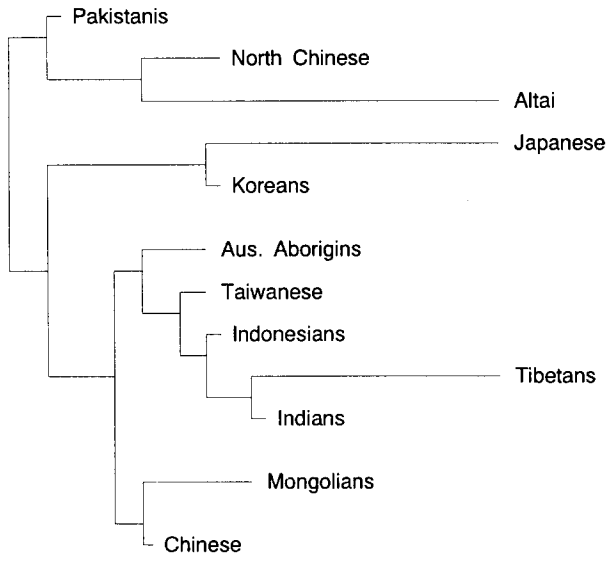

Fig. 3. Cavalli-Sforza's Chord genetic distance-NJ tree for several different ethnic groups based on the frequencies of *DYS19* and *DXYS5Y* alleles.

result could reflect a gene flow and some amount of admixture between these two populations. Also, the present survey of the *DXYS5Y* (Table 2) is consistent with the previous reports that the Y2 allele is present in samples from populations of Japan and Korea, but absent in samples representing other Asian, European, African, Oceanian, and Native American populations (Nakagome et al., 1992; Mathias et al., 1994; Hammer and Horai, 1995). This is consistent with the hypothesis that the Y2 allele tracks male lineages that originated in Korea and migrated to Japan (Hammer and Horai, 1995). The high frequency of the Y2 allele in Japanese population (~30%) could also be explained by the founder effect (M. Hammer, personal communication).

Measuring the frequencies of different Y chromosome haplotypes in different populations proves that genetically closely related populations are likely to have similar frequencies of each haplotype; more distantly related populations are expected to show progressively greater differences (Jobling and Tyler-Smith, 1995). The *DYS19* and *DXYS5Y* data presented here indicate that the genetic structure of Korean population is similar to that of the Japanese, Chinese, and Mongolian populations compared to other Asian populations (Fig. 3). Most previous reports generally agree that Koreans can be considered as a North Asian group (Nei and Roychoudhury, 1993; Cavalli-Sforza et al., 1994). The similarity for the mean haplotype diversity between Koreans and most other Asian populations suggests that the Korean population may not have a single population structure. Nei and Roychoudhury (1993) also suggested that Japanese, Koreans, Mongolians, and Tibetans shared a common ancestral population. These results would lead us to suspect that Chinese,

Table 3. A comparison of frequencies for the *DYS19/DXYS5Y* combination haplotype in the Korean population

Haplotype number	Combination haplotype	No (%) (n=480)
H1	A/Y1	21 (4.4)
H2	B/Y1	69 (14.4)
H3	C/Y1	202 (42.1)
H4	D/Y1	132 (27.5)
H5	E/Y1	34 (7.1)
H6	B/Y2	6 (1.3)
H7	C/Y2	10 (2.1)
H8	D/Y2	4 (0.8)
H9	E/Y2	2 (0.4)

Japanese, Koreans, and Mongolians shared a common ancestral population. It is also suggested that this genetic similarity supports Nei's 'out of northeast Asia hypothesis' for the origin of Japanese population (Hammer and Horai, 1995; Nei, 1995).

Hammer and Horai (1995) have reported that there are high frequencies of YAP^+ haplotype (*DYS287*) in Hokkaido/Okinawa and lower frequencies in southwestern Japan, the region closest to the Korean peninsula. An extremely low frequency of YAP^+ alleles (~1%) in the Korean population suggests a gene flow from Korea to Japan (Hammer and Horai, 1995; Kim and Ryu, 1996). In previous surveys for the mitochondrial DNA data, the closest genetic affinity between Japanese and Koreans has been found among the entire human population (Hong, 1993; Horai et al., 1996). Horai et al. (1996) have reported that ~65% of the gene pool in mainland Japanese might have migrated from the continental gene flow after the Yayoi Age. There is an evidence for a migration of Yayoi people from the Korean peninsula about 2,300 years ago to Japan (Chard, 1974). Therefore, the probability of a gene flow from Japan to Korea seems to be lower than that of the apparent large infusion of the genes with the Yayoi migration which started 2,300 years ago from Korea to Japan. Additional genetic data from a large number of informative Y chromosome markers from entire human ethnic groups are required to verify this hypothesis.

Acknowledgements

We would like to thank Dr. M. F. Hammer for help and comments during the course of this work. Special thanks go to Dr. T. Karafet for her generous assistance of phylogenetic analysis. This work was supported by a grant from the Korean Science and Engineering Foundation (KOSEF 971-0505-025-1).

References

Cann RL, Stoneking M, and Wilson AC (1987) Mitochondrial DNA and human evolution. *Nature* 325: 31-36.

Cavalli-Sforza LL and Bodmer WF (1971) The Genetics of Human Populations. W. H. Freeman & Co, San Francisco, pp 1-399.

Cavalli-Sforza LL, Menozzi P, and Piazza A (1994) The History and Geography of Human Genes. Princeton University Press, Princeton, pp 195-254.

Chard C (1974) Northeast Asia in Prehistory. University of Wisconsin Press, Madison, pp 172.
Ciminelli B, Pompei F, Malaspina P, Hammer MF, Persichetti F, Pignatti PF, Palena A, Anagnou N, Guanti G, Jodice C, Terrenato L, and Novelletto A (1995) Recurrent simple tandem repeat mutations during human Y-chromosome radiation in Caucasian subpopulations. *J Mol Evol* 41: 966-976.
Edwards A, Civitello A, Hammond HA, and Caskey CT (1991) DNA typing and genetic mapping with trimeric and tetrameric tandem repeats. *Am J Hum Genet* 11: 749-761.
Felsenstein J (1993) PHYLIP: Phylogeny Inference Package, Version 3.5 p. Joseph Felsenstein and the University of Washington, Seattle.
Gomolka M, Hundrieser J, Nürnberg P, Roewer L, Epplen JT, and Epplen C (1994) Selected di- and tetranucleotide microsatellites from chromosomes 7, 12, 14, and Y in various Eurasian populations. *Hum Genet* 93: 592-596.
Hammer MF (1994) A recent insertion of an *Alu* element on the Y chromosome is a useful marker for human population studies. *Mol Biol Evol* 11: 749-761.
Hammer MF (1995) A recent common ancestry for human Y chromosome. *Nature* 378: 376-378.
Hammer MF and Horai S (1995) Y chromosomal DNA variation and the peopling of Japan. *Am J Hum Genet* 56: 951-962.
Hammer MF, Spurdle AB, Karafet T, Bonner MR, Wood ET, Novelletto A, Malaspina P, Mitchell RJ, Horai S, Jenkins T, and Zegura SL (1997) The geographic distribution of human Y chromosome variation. *Genetics* 145: 787-805.
Hong SS (1993) The characteristics of Korean population using the genetic markers of mitochondrial and nuclear DNA. Ph.D. Thesis, Seoul National University, Korea, pp 1-220.
Horai S, Murayama K, Hayasaka K, Matsubayashi S, Hattori Y, Fucharoen G, Harihara S, Park KS, Omoto K, and Pan IH (1996) mtDNA polymorphism in East Asian populations, with special reference to the peopling of Japan. *Am J Hum Genet* 59: 579-590.
Jobling MA and Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. *Trends Genet* 11: 449-456.
Kim W and Ryu KH (1996) Analysis of polymorphism for an *Alu* element on the Y chromosome (*YAP*) in the Korean population. *Korean J Genetics* 18: 39-47.
Mathias N, Bayes M, and Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. *Hum Mol Genet* 3: 115-123.
Muller S, Gomolka M, and Walter H (1994) The Y-specific SSLP of the locus *DYS19* in four different European samples. *Hum Hered* 44: 298-300.
Nakahori Y, Tamura T, Yamada M, and Nakagome Y (1989) Two 47z [*DXYS5*] RFLPs on the X and Y chromosome. *Nucleic Acids Res* 17: 2152.
Nakagome Y, Young SR, Akane A, Numabe H, Jin DK, Yamori Y, Seki S, Tamura T, Nagafuchi S, Shiono H, and Nakahori Y (1992) A Y-associated allele may be characteristic of certain ethnic groups in Asia. *Ann Hum Genet* 56: 311-314.
Nei M (1987) Molecular Evolutionary Genetics. Columbia University Press, New York, pp 179.
Nei M (1995) The origins of human populations: genetic, linguistic and archeological data. In: Brenner S and Hanihara K (eds), The Original Past of Modern Humans as Viewed from DNA. World Scientific, Singapore, pp 71-91.
Nei M and Roychoudhury AK (1993) Evolutionary relationships of human populations on a global scale. *Mol Biol Evol* 10: 927-943.
Ngo KY, Vernaud G, Johnsson C, Lucotte G and Weissenbach J (1986) A DNA probe detecting multiple haplotypes of the human Y chromosome. *Am J Hum Genet* 38: 407-418.
Richards B, Skoletsky J, Shuber AP, Balfour R, Stern RC, Dorkin HL, Parad RB, Witt D, and Klinger KW (1993) Multiplex PCR amplification from the *CFTR* gene using DNA prepared from buccal brushes/swabs. *Hum Mol Genet* 2: 159-163.
Roewer L, Arnemann J, Spurr NK, Grzeschik KH, and Epplen JT (1992) Simple repeat sequences on the human Y chromosome are equally polymorphic as their autosomal counterparts. *Hum Genet* 89: 389-394.
Saitou N and Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406-425.
Santos FR, Pena SDJ, and Epplen JT (1993) Genetic and population study of a Y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. *Hum Genet* 90: 655-656.
Santos FR, Gerelsaikhan T, Munkhtuja B, Oyunsuren T, Epplen JT, and Pena SDJ (1996) Geographic differences in the allele frequencies of the human Y-linked tetranucleotide polymorphism *DYS19*. *Hum Genet* 97: 309-313.
Spurdle AB, Hammer MF, and Jenkins T (1994) The Y *Alu* polymorphism in southern African populations and its relationship to other Y-specific polymorphisms. *Am J Hum Genet* 54: 319-330.

[Received March 26, 1998; accepted April 24, 1998]