J Hum Genet (1999) 44:240–245

ORIGINAL ARTICLE

**Toshikatsu Shinka · Keiko Tomita · Tatsushi Toda
Svetlana E. Kotliarova · Juwon Lee · Yoko Kuroki
Dong Kyu Jin · Katsushi Tokunaga · Hideki Nakamura
Yutaka Nakahori**

Genetic variations on the Y chromosome in the Japanese population and implications for modern human Y chromosome lineage

Received: February 25, 1999 / Accepted: April 5, 1999

Abstract A polymorphism in the coding sequence of the *SRY* gene was found by single-strand conformation polymorphism (SSCP) and direct sequencing analysis. The new allele of the *SRY* gene, which is raised by a C-to-T transition in the 155th codon, was found in 24% of Honshu, 35% of Okinawan, and 51% of Korean males respectively, whereas it was not observed among 16 Caucasian and 18 Negroid males. A haplotype analysis of the Y chromosome was carried out in Japanese, Korean, Caucasian and Negroid populations, using a combination of the polymorphisms in *SRY*, DXYS5Y, DYS287, and DXYS241Y loci. The results indicated that the Y chromosomes can be classified into seven haplotypes (Ia, Ib, Ic, IIa, IIb, III, IV). However, of these seven, only four (Ia, IIa, III, IV) were observed in the Japanese population. Furthermore, the presumed haplotype C, Y1, YAP, $(CA)_{14}$, from which haplotype III was probably derived, was not found in any populations in this study. The regional distribution of each haplotype revealed that type III is more frequently observed in Okinawa (16%) and in Korea (21%) than in Honshu (4.4%). The haplotype analysis of the Y chromosome may contribute to the exploration of the origin of Japanese and the relationship between east Asian populations.

T. Shinka · J. Lee · Y. Kuroki · H. Nakamura · Y. Nakahori (✉)
Department of Public Health, School of Medicine, University of Tokushima, Kuramoto-cho, Tokushima 770-0085, Japan
Tel. +81-886-33-7075 (ext. 2272); Fax +81-886-33-7453
e-mail: nakahori@basic.med.tokushima-u.ac.jp

K. Tomita · S.E. Kotliarova · Y. Kuroki · K. Tokunaga
Department of Human Genetics, Graduate School of International Health, University of Tokyo, Tokyo, Japan

T. Toda
Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan

D.K. Jin
Department of Pediatrics, Samsung Medical Center, Seoul, Korea

Key words Y chromosome · Polymorphism · Haplotype · SRY · DXYS5Y · YAP · DXYS241Y · Japanese

Introduction

The Y chromosome has unique characteristics because it is a single haploid unit in the human genome that is passed only from father to son. Thus it represents the patrilineal contribution to the male genome. DNA markers residing in the non-recombining portion of the human Y chromosome were shown to be useful for tracing male-specific gene flow and also in human evolution studies (Poobo 1995; Dorit et al. 1995; Jobling and Tyler-Smith 1995; Hammer 1995; Whitfield et al. 1995; Thomas et al. 1998).

Most of the Y chromosome polymorphisms are known to be specific to populations (Nakahori et al. 1989; Torroni et al. 1994; Seielstad et al. 1994; Jobling et al. 1994; Mathias et al. 1994; Pena et al. 1995; Underhill 1996). The YAP (Y chromosome Alu insertion polymorphism) at locus DYS287 is found among sub-Saharan African, North African, and European populations, while most Asian populations other than the Japanese lack the YAP element (Spurdle et al. 1994; Hammer and Horai 1995). On the other hand, the 47z/*Stu*I polymorphism at locus DXYS5Y (Nakahori et al. 1989) was found in Japan, Korea, and Taiwan, but was absent in other populations (Nakagome et al. 1992; Lin et al. 1994). DXYS241Y is a polymorphic CA dinucleotide repeat located on the X-Y homologous region shown to be useful for analysis of the human Y chromosome lineage (Kotliarova et al. 1999).

During a study of patients with abnormal sex differentiation, we found a novel polymorphism in the *SRY* (sex determining region Y chromosome) gene. Here we report the polymerase chain reaction (PCR)-based detection of this polymorphism, as well as the newly developed 47z/*Stu*I PCR systems. A haplotype analysis, including *SRY*, 47z/*Stu*I, and YAP allowed us to set the order of the appearance of these polymorphisms. The origin of Japanese males and evolution of the Y chromosome are discussed.

Materials and methods

DNA extraction. Genomic DNAs were prepared from peripheral leukocytes according to the standard method (Maniatis et al. 1989). The samples from Japanese were collected from volunteer donors in Honshu (mainly from Tokyo) and Okinawa in Japan. The blood samples of Caucasians and Negroids were previously described elsewhere (Nakagome et al. 1992).

PCR conditions. The reaction was performed in a total volume of 20 μl containing 66 ng genomic DNA, 1.5 mM $MgCl_2$, 10 mM TrisHCl-buffer (pH8.3), 50 mM KCl, 0.001% (w/v) gelatin, 0.25 mM of each dNTP, 1 μM of each primer, and 0.1 U Ampli Taq DNA polymerase. It consisted of a cycle of 94°C for 3 min, 55°C for 2 min, and 72°C for 2 min, followed by 25 cycles of 94°C for 1 min, 55°C for 2 min, and 72°C for 2 min, and a final extension at 72°C for 12 min.

The YAP element and the polymorphism on the DXYS241Y locus were detected using a PCR technique according to the methods described by Hammer and Horai (1995) and Kotliarova et al. (1999), respectively.

Single-strand conformation polymorphism (SSCP) analysis. SSCP was performed in a total volume of 5 μl containing 33 ng genomic DNA, 1.5 mM $MgCl_2$, 10 mM TrisHCl-buffer (pH8.3), 50 mM KCl, 0.001% (w/v) gelatin, 0.25 mM of each dNTP, 1 μM 32P-dATP (Amersham, Arlington, Heights, IL, USA), 1 μM of each primer, and 0.1 U Ampli Taq DNA polymerase (Perkin-Elmer, Norwalk, CT, USA). The reaction consisted of a cycle of 94°C for 3 min, 65°C for 2 min, and 72°C for 2 min, followed by 25 cycles of 94°C for 1 min, 65°C for 2 min, and 72°C for 2 min, and a final extension at 72°C for 12 min. Five μl of formamide dye (95% formamide, 20 mM ethylenediaminetetraacetic acid (EDTA), 0.05% bromophenol blue, 0.05% xylene cyanol) was added to 5 μl of radioactive PCR products. After denaturing the PCR products at 95°C for 5 min, we loaded them (3 μl) onto 6% acrylamide gel containing 5% glycerol. Electrophoresis was carried out at room temperature for 2.5 h at 35–40 W in 1 × TBE (45 mM Tris-Borate, 1 mM EDTA). The gels were dried and exposed to X-ray films overnight.

Direct sequencing analysis of SRY. Single-stranded template DNA was prepared using a biotinylated primer, as described (Hultman et al. 1989; Iida et al. 1994). The dideoxy termination method was carried out, using the Sequenase-sequencing kit (USB) and S^{35}-dATP (Amersham).

Results

SRY polymorphism. Since the primers we have been using for the screening of the SRY mutation produced little shift of a band, other primers were designed, according to the nucleotide sequence reported by Su et al. (Su et al. 1993). When we used the primers SRY13: 5′-GCCGAAGAATTGCAGTTTGC and SRY14: 5′-GTTGATGGGCGGTAAGTGGC, the two alleles resolved well. PCR-SSCP analysis of the SRY is shown in Fig. 1A. The direct sequence analysis revealed a synonymous mutation raised by a C-to-T transition at the 155th codon serine from the initiation codon methionine. This substitution abolishes the site of *Bso*FI endonuclease. To detect the difference of the *Bso*FI restriction site, the *SRY* gene was amplified by the primers SRY13 and SRY14 and digested with *Bso*FI (Fig. 1B). The allele harboring the same sequence as that published (Su et al. 1993) was designated as the C allele, and the other allele as the S allele. The S allele was present in 89 (32%) of the 277 Japanese and 20 (51%) of the 39 Korean males, respectively, while it was absent in both the 16 Caucasian and the 18 Negroid males. The frequency of the S allele was not significantly different between Honshu (26%) and Okinawa (35%) males (χ^2 test; $P = 0.1$).

*System for 47z/Stu*I *polymorphism.* The probe p47z detects polymorphisms on the short arm of the Y chromosome and long arm of the X chromosome (Nakahori et al. 1989). When genomic DNAs were digested with *Stu*I, the Y chromosome polymorphism was detected by Southern hybridization, and either a 17-kb (Y1) or a 5.3-kb (Y2) band was observed. The surrounding sequence of the *Stu*I polymorphic site was determined from the cosmid clone positive for the probe. Since the polymorphic site was in an Alu repeat sequence, we tried two pairs of primers for a long product and a short product. Both of them worked well. The primers for the long product are NS1-F: 5′-TTTGCTTCTCATTTCATCTG and NS1-R: 5′-TTAGATGAATTGTTGTGCTG. The primers for the short product are NS2F: 5′-TGAGTCAATGTCAATGAATC and NS2R: 5′-TAGTTACGCCTTGGCATAAC. The alleles were separated by the digestion with *Stu*I (Fig. 2).

The observed frequency of the Y2 allele in Honshu males (21.1%) was similar to that in Okinawans (19.8%) (χ^2 test; $P = 0.8$). In Koreans, it was higher (31%) than previously reported (Nakagome et al. 1992).

Analysis of YAP. The YAP polymorphism was detected in 31% (28) of Honshu and 34% (64) of Okinawan males. These frequencies were lower than those previously reported by Hammer (Hammer and Horai 1995). In the Caucasian and Negroid males, the YAP elements were observed at frequencies of 72% and 25%, respectively.

*Haplotype analysis of the SRY, 47z/*Stu*I, YAP, and DXYS241Y polymorphisms.* The haplotypes of 90 Y chromosomes in Honshu and 187 chromosomes in Okinawan males were constructed using a combination of SRY, 47z/*Stu*I, YAP, and DXYS241Y polymorphisms. The result of the haplotype analysis is summarized in Table 1. The Y chromosomes possessing the C allele always have the Y1 allele and can be divided into two groups based on the YAP (haplotypes I and II). Furthermore, haplotype I can be divided into three subtypes (Ia, Ib, and Ic), and haplotype II

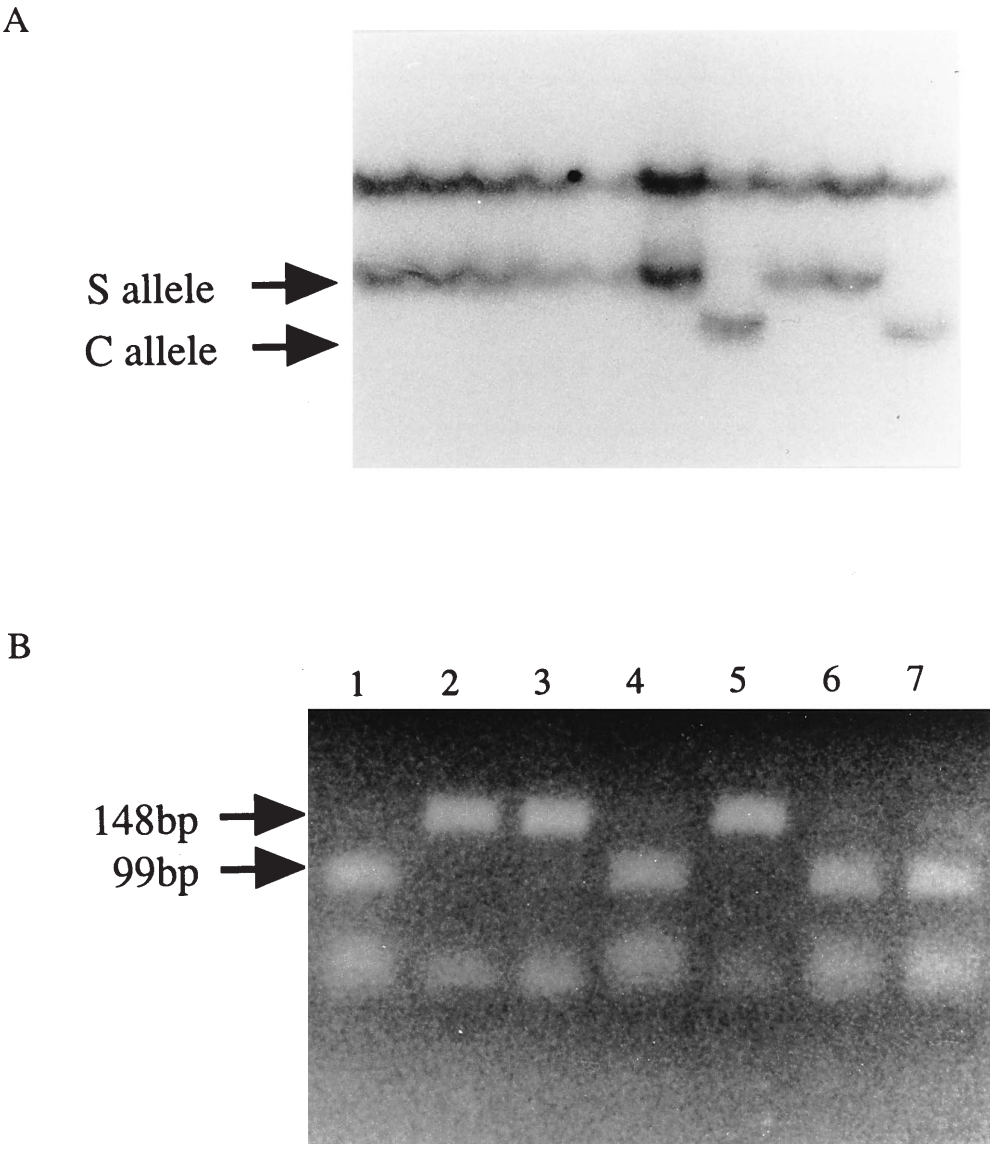

Fig. 1. **A** Single-strand conformation polymorphism (SSCP) analysis of the SRY polymorphism using primers 13 and 14. The *upper bands* show the S alleles, and the *lower ones* the C alleles. **B** Polymerase chain reaction – restriction fragment length polymorphism (PCR-RFLP) analysis of the polymorphism using *Bso*FI endonuclease. The C allele has a *Bso*FI site and gives a 99 bp band. The S allele does not have a *Bso*FI site and gives a 148 bp band. *Lanes 1, 4, 6, and 7* show the C alleles and *lanes 2, 3, and 5* the S alleles

into two (IIa and IIb) based on the polymorphism of the DXYS241Y locus, respectively. The Y chromosomes with the S allele are always YAP- and $(CA)_{14}$, while they can be divided into two groups based on the 47z/*Stu*I polymorphism (haplotypes III and IV). The constant association of the Y2 allele with the YAP- allele is consistent with the observations of Hammer et al. (Hammer and Horai 1995). The frequency of haplotype III was significantly higher in the Okinawan than in the Honshu population (χ^2 test; $P = 0.008$).

Furthermore, some blood samples derived from Koreans, Caucasians, and Negroids were also analyzed (Table 1). Although haplotypes III and IV were not observed in any samples from Caucasians and Negroids, they were detected in 51% of the blood samples from Koreans.

Regional distribution of the haplotypes on the Y chromosome in Okinawa. The frequencies of each haplotype were determined in Okinawa, including the areas of Itoman, Yomitan, Katsuren, Gushikami, Iriomote, and Hateruma

Table 1. Association among polymorphisms of the *SRY* gene, YAP, 47z/*Stu*I, and DXYS241Y in different populations

	Haplotype							
	Ia	Ib	Ic	IIa	IIb	III	IV	
Locus								
SRY	C	C	C	C	C	S	S	
47z/*Stu*I	Y1	Y1	Y1	Y1	Y1	Y2	Y2	
YAP[a]	−	−	−	+	+	−	−	
DXYS241Y[b]	15	16	17	15	16	14	14	
Population	No. of chromosomes							Total
Honshu	39	0	0	28	0	4	19	90
Okinawa	57	0	0	64	0	29	37	187
Korean	18	1	0	0	0	8	12	39
Caucasian	8	2	2	4	0	0	0	16
Negroid	3	0	2	12	1	0	0	18

[a] +, Alu element is present; −, Alu element is absent
[b] Number of repeats

Table 2. The frequencies of each haplotype in Okinawan males

	Haplotype							
	Ia	Ib	Ic	IIa	IIb	III	IV	
Population	No. of chromosomes (%)							Total
Honshu	39 (43.3)	0	0	28 (30.1)	0	4 (4.4)	19 (21.1)	90 (100)
Okinawa[a]	57 (30.5)	0	0	64 (34.2)	0	29 (15.5)	37 (19.8)	187 (100)
Itoman	22 (27.5)	0	0	38 (47.5)	0	8 (10.0)	12 (15.0)	80 (100)
Yomitan	9 (28.1)	0	0	10 (31.3)	0	7 (21.9)	6 (18.8)	32 (100)
Katsuren	10 (34.5)	0	0	8 (27.6)	0	5 (17.2)	6 (20.7)	29 (100)
Gushikami	8 (42.1)	0	0	7 (36.8)	0	1 (5.3)	3 (15.8)	19 (100)
Iriomote	4 (20.0)	0	0	1 (5.0)	0	8 (40.0)	7 (35.0)	20 (100)
Hateruma	4 (57.1)	0	0	0 (0)	0	0 (0)	3 (42.9)	7 (100)

[a] Okinawa includes Itoman, Yomitan, Katsuren, Gushikami, Iriomote, and Hateruma

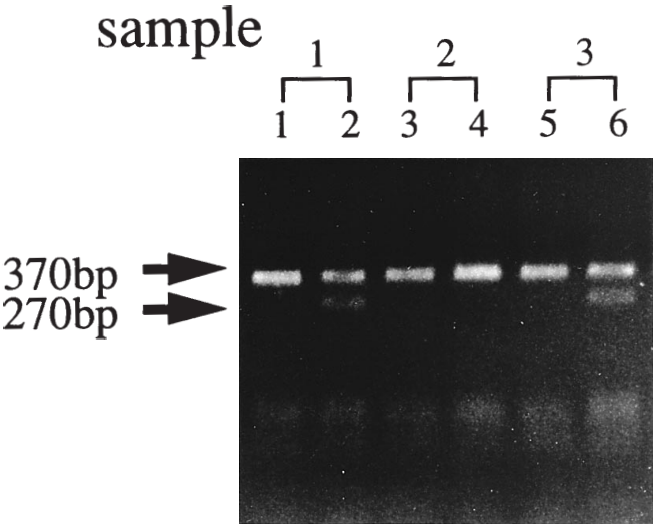

Fig. 2. Detection of the 47z/*Stu*I polymorphism using primers NS2-F and NS2-R. *Odd lanes* indicate undigested PCR products, and *even lanes* digested products. The common 370 bp products are derived from both the X chromosome and the Y chromosome. However, in the individuals with the Y2 allele (*samples 1 and 3*), the products were digested with *Stu*I, which resulted in short fragments

(Table 2). Although haplotype III was observed at a very low frequency in Honshu (4.4%), it was more frequent in Yomitan (22%), Katsuren (17%), and Iriomote (40%).

Discussion

We have found a novel polymorphism in the *SRY* gene and developed a PCR-based detection system for it, as well as the 47z/*Stu*I polymorphism. With these and the polymorphisms of YAP and DXYS241, a novel $(CA)_n$ X-Y homologous locus, we constructed the haplotype of the Y chromosome in different populations and classified the Y chromosome into seven types.

To date, some polymorphisms of the *SRY* gene have been reported to locate in the non-coding region (Jobling and Tyler-Smith 1995; Hurles et al. 1998). On the other hand, the newly found polymorphism resides in the coding region of the *SRY* gene. In the nucleotide sequences of the chimpanzee, gorilla, and orangutan, the code for the 155th codon serine is AGC (Whitfield et al. 1993). Therefore, the S allele with the codon AGT may be the result of a recent substitution.

Questions arise as to the order of the appearance of the polymorphisms. The fact that the Y2 allele is always associated with the S and YAP− alleles indicates that it arose on the Y chromosome having both of these alleles. Since the S allele is always associated with YAP−, but YAP− is not always associated with the S allele, the S allele may have arisen on the Y chromosome having the YAP− allele. Hammer hypothesized that a YAP insertion occurred once in the evolution (Hammer 1994; Hammer and Horai 1995; Hammer 1995). If this is the case, we can assume that the original haplotype was C,Y1,YAP− (haplotype I) and the branching of the YAP+ (haplotype II) and S allele (haplotype III) happened independently in different Y chromosomes. The Y chromosomes with S,Y2,YAP− (haplotype IV) alleles were branched off from haplotype III (Fig. 3).

Based on the constructed haplotypes, we determined the frequency of each haplotype in Japanese, Caucasian, Negroid, and Korean populations. The Y chromosomes with YAP (haplotype IIa) were observed in 34.2% of Okinawans. This frequency was different from the one previously reported by Hammer and Horai (1995). However, even within Okinawa, the frequency of the Y chromosome with YAP ranges from 47.5% in Iotman to 0% in Hateruma. So the area for the blood sampling in Okinawa may be different between ours and Hammer's. Okinawa is part of the route across the East China sea and consists of many islands. Therefore, we may assume gene flow from China, Taiwan, Japan, and Korea into the islands. In any event, it is necessary to analyze larger numbers of samples to determine accurately the frequencies of each haplotype of the Y chromosome in Okinawa.

Although we found no Y chromosomes with YAP+ in our Korean samples, Y chromosomes harboring YAP+ in a very small population of Koreans were recently reported (Kim et al. 1998).

Reports have revealed that haplotype analysis of the Y chromosome is able to provide clues as to the origin of some

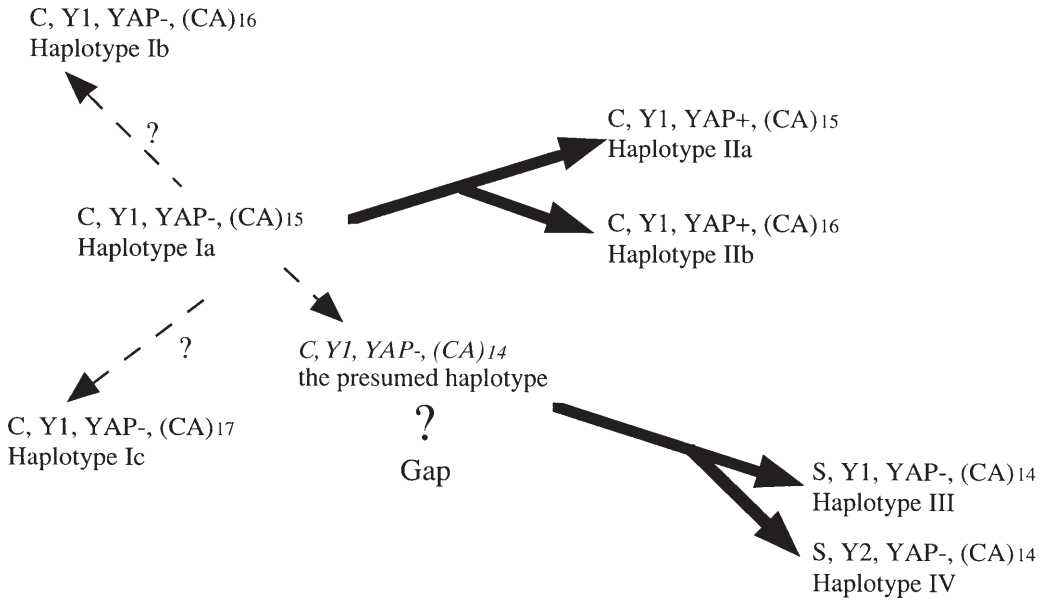

Fig. 3. Schematic representation of the relationship among the Y chromosome haplotypes based on the polymorphisms in *SRY*, 47z/*Stu*I, YAP, and DXYS241Y. The haplotypes observed were linked by the rationale described in the text. Haplotypes Ib (C, Y1, YAP−, $(CA)_{16}$), and Ic (C, Y1, YAP−, $(CA)_{17}$), which were not observed in the Japanese population, are indicated with the *dotted arrows* in this schema because the relationship among haplotypes Ia, Ib, and Ic is not clear. The *italic* letters show the presumed haplotype which was not observed in this study

historically important family and human populations (Bamshad et al. 1998; Thomas et al. 1998). The history of the Y chromosome is different from that of mt DNAs, reflecting social histories, such as mating structures and the differential behavior of males and females in migrations, wars, and colonizations (Bamshad et al. 1998; Jobling and Tyler-Smith 1995). The geological distribution of these polymorphisms is a matter of interest, since the era in which each polymorphism is generated differs.

We have previously reported that the distribution of the Y2 allele is restricted to certain regions of east Asia, such as Japan, Taiwan, and Korea (Nakagome et al. 1992; Lin et al. 1993). As the Y2 allele is always associated with the S allele, it would be worth exploring the distribution of the S allele. Interestingly, the possible haplotype C,Y1,YAP−, $(CA)_{14}$, from which haplotype III is presumably derived, was not observed in any samples in this study. Therefore, studying the distribution of the S allele and the presumed haplotype C,Y1,YAP−, $(CA)_{14}$ in Asia or other regions of the world may provide a clue as to the origin of the east Asian people, including the Japanese.

We carried out a haplotype analysis of the Y chromosome in a Japanese population by a newly developed PCR-based analysis of two polymorphisms on the Y chromosome. The construction of haplotypes using these PCR-based polymorphisms will be useful for pursuing the origin of the Japanese and the relationship between east Asian populations, and will contribute to the evolutionary tree of the modern human Y chromosome.

Acknowledgments This work was supported by a Grant-in-Aid for General Scientific Research from the Ministry of Education Science, Sports, and Culture of Japan.

References

Bamshad MJ, Watkins WS, Dixon ME, Jorde LB, Rao BB, Naidu JM, Prasad BVR, Rasanayagam A, Hammer MF (1998) Female gene flow stratifies Hindu castes. Nature 395:199–652

Dorit RL, Akashi H, Gilbert W (1995) Absence of polymorphism at the ZFY locus on the human Y chromosome. Science 268:1183–1185

Hammer MF (1994) A recent insertion of 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, Horai S (1995) Y Chromosomal DNA variation and the peopling of Japan. Am J Hum Genet 56:951–962

Hultman T, Stahl S, Hornes E, Uhlen M (1989) Direct solid phase sequencing of genomic and plasmid DNA using magnetic beads as solid support. Nucleic Acids Res 17:4937–4946

Hurles ME, Irven C, Nicholson J, Taylor PG, Santos FR, Loughlin J, Jobling MA, Sykes BC (1998) European Y-chromosome lineages in Polynesians: a contrast to the population structure revealed by mtDNA. Am J Hum Genet 63:1793–1806

Iida T, Nakahori Y, Komaki R, Mori E, Hayashi N, Tstsumi O, Taketani Y, Nakagome (1994) A novel nonsense mutation in the HMG box of the SRY gene in a patient with XY sex reversal. Hum Mol Genet 3:1437–1438

Jobling MA (1994) A survey of long-range DNA polymorphisms on the human Y chromosome. Hum Mol Genet 3:107–114

Jobling MA, Tyler-Smith C (1995) Fathers and sons: the Y chromosome and human evolution. Trends Genet 11:449–456

Kim W, Shin DJ, You Y, Kim J (1998) Y-Specific DNA polymorphisms of the *YAP* element and the locus DYS19 in the Korean population. J Hum Genet 43:195–198

Kotliarova ES, Toda T, Takenaka O, Matsushita I, Hida I, Shinka T, Goto J, Tokunaga K, Nakagome Y, Nakahori Y (1999) A novel $(CA)_n$ marker DXYS241 on the non-recombinant part of the human Y chromosome. Hum Biol 71:263–277

Lin SJ, Tanaka K, Lenard W, Gerelsaikhan T, Dashnyam B, Nyamkhishig S, Hida A, Nakahori Y, Omoto K, Crawford MH, Nakagome Y. (1994) A Y-associated allele is shared among a few ethnic groups of Asia. Jpn J Hum Genet 39:299–304

Maniatis T, Fritsch E, Sambrook J (1989) Molecular cloning: a laboratory manual. Second edn. Cold Spring Harbor, New York.

Mathias N, Bayes M, Tyler-Smith C (1994) Highly informative compound haplotypes for the human Y chromosome. Hum Mol Genet 3:115–123

Nakagome Y, Young SR, Akane A, Numabe H, Jin DK, Amori YY, Seki S, Tamura T, Nagafuchi S, SHIONOH, Nakahori Y (1992) A Y-associated allele may be characteristic of certain ethnic groups in Asia. Ann Hum Genet 56:311–314

Nakahori Y, Tamura T, Yamada M, Nakagome Y (1989) Two 47Z [DXYS5] RFLPs on the X and the Y chromosome. Nucleic Acids Res 17:2152

Pena SDJ, Santos FR, Bianchi NO, Bravi CM, Carnese FR, Rothhammer F, Gerelsaikhan T, Munkhtuja B, Oyunsuren T. (1995) A major founder Y-chromosome haplotype in Amerindians. Nature Genet 11:15–16

Poobo S (1995) The Y chromosome and the origin of all of us (men). Science 268:1141–1142

Seielstad MT, Herbert JM, Lin AA, Undehill PA, Ibrahim M, Vollrath Dand, Cavalli-Sforza LL (1994) Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition. Hum Mol Genet 3:2159–2161

Spurdle AB, Hammer MF, 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

Su H, Lau YC (1993) Identification of the transcriptional unit, structural organization, and promoter sequence of the human sex-determining region Y (SRY) gene, using a reverse genetic approach. Am J Hum Genet 52:24–38

Thomas MG, Ben-Ami H, Skorecki K, Parfitt T, Bradman N, Goldstein DB (1998) Origins of Old Testament priests. Nature 394:138–140

Torroni A, Chen Y-S, Semino O, Santachiara-Benerecetti AS, Scott CR, Lott MT, Winter M, Wallace DC (1994) Mitochondorial DNA and Y-chromosome polymorphisms in four Native American populations from south Mexico. Am J Hum Genet 54:303–318

Underhill PA, Jin L, Zemans R, Oefner PJ, Cavalli-Sforza (1996) A pre-Colombian Y chromosome-specific transition and its implications for human evolutionary history. Proc Natl Acad Sci USA 93:196–200

Whitfield LS, Lovell-Badge R, Goodfellow PN (1993) Rapid sequence evolution of the mammalian sex-determining gene SRY. Nature 364:713–715

Whitfield LS, Sulston JE, Goodfellow PN (1995) Sequence variation of the human Y chromosome. Nature 378:379–380