

HYPERVARIABLE POLYMORPHISM OF AUTOSOMAL ORIGIN DETECTED BY THE Y-CHROMOSOME DERIVED PROBE, pHY10

Tsuyoshi YOKOI,¹ Kaoru SAGISAKA,¹ Yutaka NAKAHORI,²
and Yasuo NAKAGOME²

¹*Department of Forensic Medicine, Tohoku University School of Medicine,
Sendai 980, Japan*

²*Department of Congenital Anomalities Research, National Children's
Medical Research Center, Setagaya-ku, Tokyo 154, Japan*

Summary A recombinant DNA probe (pHY10) hybridizing specifically to human DNA family DYZ1, 3,000 copies of which are present on the long arm of the Y chromosome, was used for probing human genome DNA digested with various restriction enzymes. To our surprise, the probe detected a hypervariable polymorphism of autosomal origin in human DNA when digested with *TaqI*. None of other 12 restriction enzymes revealed polymorphic patterns. Codominant segregation of the polymorphism was established in family studies. This probe has been widely used in the detection of the Y chromosome. Its ease of availability as well as highly discriminating polymorphic pattern makes it potentially very useful for forensic and human genetic purposes.

Key Words hypervariable polymorphism, Y-chromosome, RFLP, DYZ1

INTRODUCTION

A Y-specific DNA family, DYZ1, was first described by Cooke (1976) as a 3.4 kb band in *HaeIII* digests of human male DNA. It was later shown that this repeated DNA family was a major component of the distal part of Yq (Yq=the long arm of a Y) (Bostock *et al.*, 1978). The DYZ1 family contains *EcoRI*-3564 bp fragments (pHY10) composed of a tandem array of pentanucleotides with a few insertions/deletions. Two hundred and twenty-nine (32.1%) of the pentanucleotides were TTCCA, 297 (41.7%) were single-base substituents of them and 139 (19.5%) were two-base substituents (Nakahori *et al.*, 1986). The pHY10 has 64 *TaqI* sites and generates very small fragments (av.=56 bp) when digested with *TaqI*.

Received June 2, 1989; revised version received August 4, 1989; Accepted August 17, 1989

Southern-blot patterns of *EcoRI* (or *HaeIII*)-digested male and female DNA hybridized with the pHY10 reveal a male specific 3.56 kb band. The families of both satellite II and III also contain DNA fragments with a TTCCA motif within them (Prosser *et al.*, 1986).

In the present study, male and female DNA was digested with 13 different restriction enzymes and probed by pHY10. The pHY10-*TaqI* combination revealed, hypervariable polymorphic bands in both males and females of a complexity comparable to a minisatellite (Jeffreys *et al.*, 1986).

MATERIALS AND METHODS

Preparation of samples. To prepare DNA from peripheral blood, buffy coat was mixed with several aliquots of 0.2% NaCl containing 10 mM ethylenediamine tetraacetic acid (EDTA) at 4°C. DNA was extracted by the standard technique (Maniatis *et al.*, 1982). Blood and semen stains were prepared on cotton fabric and aged at room temperature for one month. Two μg of DNA was obtained from 8 cm² of blood stain or 4 cm² of semen stain. The fabric was shredded for extraction.

Labeling of probes. Standard procedures were used for bacterial transformation and plasmid DNA isolation (Maniatis *et al.*, 1982). Y specific probe (pHY10) was labeled with [³²P]dCTP (3,000 $\mu\text{Ci}/\text{mmol}$) using a Multiprime DNA labeling system (Amersham).

Southern hybridization. Thirteen different restriction enzymes, (REs; *EcoRI*, *HaeIII*, *TaqI*, *HinfI*, *ClaI*, *AsuII*, *PstI*, *XbaI*, *HindIII*, *RsaI*, *MspI*, *AluI* and *AccI* (Boehringer Mannheim, G.F.R.), were used for the digestion of DNA specimens and incubated at either 37°C (REs other than *TaqI*) or 65°C (*TaqI*) for 2–3 hr using 10–12 total units of RE per sample. The digested DNA fragments were electrophoresed in a 0.8% agarose gel. The DNA fragments were then transferred to a nitrocellulose membrane.

Prehybridization was carried out at 42°C in a solution containing 50% formamide, 5 × SSPE (1 × SSPE = 0.15 M sodium chloride, 0.01 M sodium dihydrogen phosphate 1 mM EDTA), 0.2 mg/ml salmon sperm DNA, 0.1% SDS, and 5 × Denhardt's solution [1 × Denhardt's solution = 0.2 g/liter Ficoll (Sigma, type 400), 0.2 g/liter polyvinylpyrrolidone (PVP, Kodak), and 0.2 g/liter bovine serum albumin (Sigma, type V)]. The filters were hybridized with probes at 42°C for 12 hr in the same solution as prehybridization except that the concentration of salmon sperm DNA was reduced by half. The filters were then washed twice for 10 min at room temperature with a solution containing 2 × SSPE plus 0.1% SDS and then with a solution of 0.1 × SSPE for a few minutes. The filters were then air dried. Autoradiography was carried out at –90°C using Fuji X-ray film for one day. Densitometry of X-ray films was carried out using a CS-9000 densitometer (Shimadzu Seisakusho Co., Ltd.).

RESULTS

Southern blots of *TaqI* digested DNA are shown in Fig. 1. DNA samples studied: unrelated individuals, paired DNA samples of leukocytes and semen from the same individuals, a monozygotic and a dizygotic twin pair and a family (parents and a child, see legend of Fig. 1). Samples include both male and female DNA. Blood and semen stains were aged for one month and compared with fresh samples.

About 15–20 bands are identifiable in individual lanes. The monozygotic

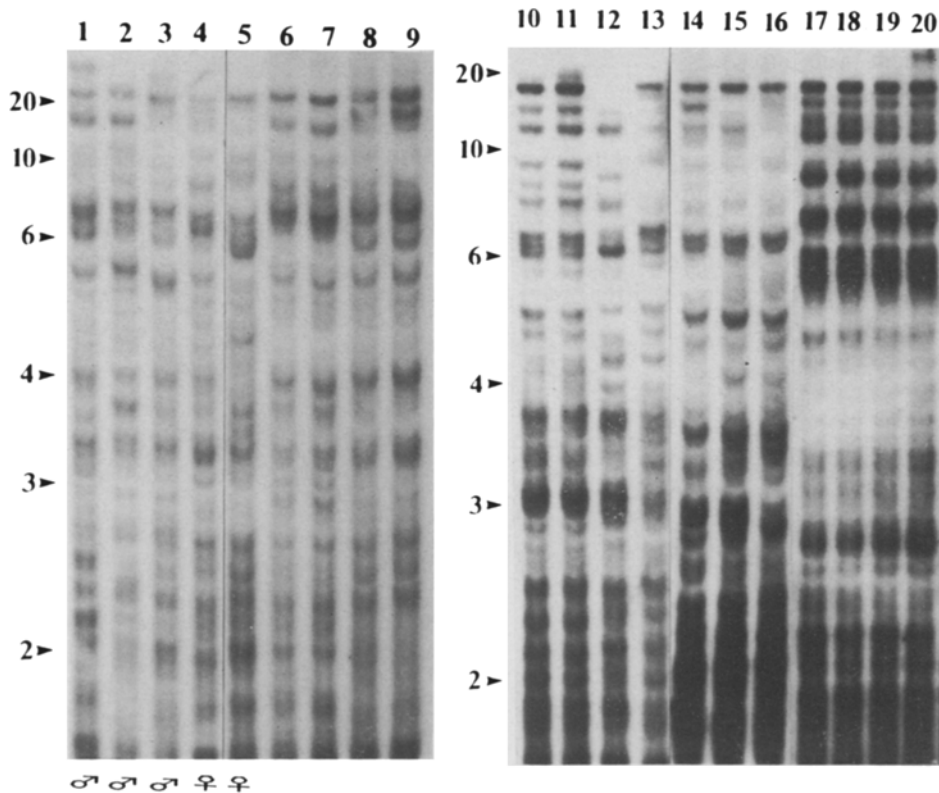


Fig. 1. Southern blots of *TaqI* digested DNA are shown. Probe was ^{32}P -labeled pHY10. Lanes 1–5 are unrelated individuals. Lanes 1–3 are male and 4–5 female. Lanes 6–9 are 2 pairs of DNA samples from 2 individuals, leukocytes (lanes 6 and 8) and semen (lanes 7 and 9). Lanes 10 and 11 represent a monozygotic twin pair. Lanes 12 and 13 are from a dizygotic twin pair (zygosity diagnosis was based on blood-group and isozyme markers). Lanes 14, 16 and 15 are from a family, a father, a mother and their son. Lanes 17 to 20 show from the same individual, leukocytes (lane 17), one-month-old blood stain (lane 18), semen (lane 19) and one-month-old semen stain (lane 20). The arrowheads indicate the approximate size of the bands.

twins as well as two pairs of leukocyte-semen samples showed identical patterns. Unrelated individuals showed patterns distinguishable from one another. One month-old blood and semen stains do not seem to show recognizable changes in patterns.

Digestion with *Cla*I (ATCGAT) and *Asu*II (TTCGAA), each of which has a *Taq*I (TCGA) site within it, did not show any polymorphism. Most DNA appeared to remain over 20 kb. Our previous work has shown that many *Taq*I sites overlap with *Hinf*I sites on pHY10 (Nakahori *et al.*, 1986). It was thus expected that *Hinf*I also would produce polymorphic bands. No clear bands were observed in the 2.0–30.0 kb range although some poorly resolved bands were detected below 2.0 kb. None of the other 9 enzymes revealed distinct polymorphic bands by the present procedures.

*Taq*I digested bands were traced using a densitometer. Examples of the tracings are shown in Fig. 2. Results are also summarized in Table 1 which includes two familial cases in addition to the one shown in Fig. 1. Results from lanes 17–20 in the Fig. 1 were omitted as they were intended to compare fresh and one-month old blood/semen stains. Gene frequencies were calculated based on examinations of 95 unrelated Japanese (52 males and 43 females). Twenty families composed of both parents and their child(ren) were investigated.

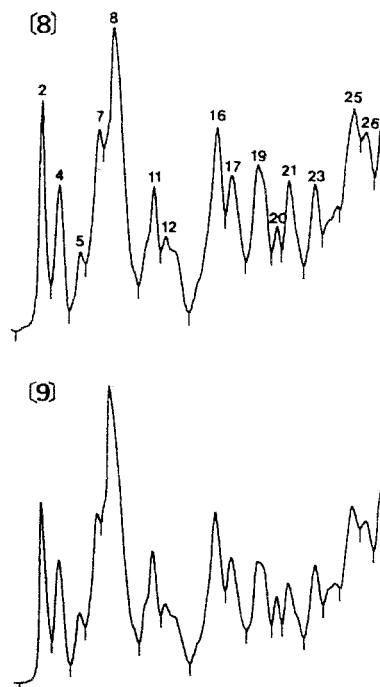


Fig. 2. Densitometric tracing of lanes 8 and 9 in Fig. 1. The first tracing is from DNA of leukocytes and the second is from the semen of the same individual. Both tracings are quite similar.

Table 1. Summary of densitometric tracing of Southern blots.

Lanes 1 to 16 correspond to those of Fig. 1. Lanes 17 to 20 were omitted since they resulted from the comparison of fresh and one-month-old stain samples. Lanes 21-24 and 25-27 are from two families. Lanes 21 and 25 are from fathers, lanes 24 and 27 are from mothers, lanes 22 and 23 are from two daughters of one of the parents (21 and 24) and lane 26 is from a son of other parents (25 and 27). Peaks are serially numbered starting from the largest. The approximate size of individual peaks are shown in kb. + or - represents the presence or absence of a particular band in a particular lane. Gene frequencies are based on 95 unrelated individuals.

Peaks Sizes	Lanes																Gene frequency										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		21	22	23	24	25	26	27			
1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.07	
2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.49
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.65
4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.22
5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.79
6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.45
7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.63
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.37
10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.31
11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.83
12	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.71
13	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.45
14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.60
15	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.44
16	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.45
17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.68
18	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.65
19	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.83
20	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.62
21	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.70
22	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.55
23	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
24	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.82
25	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
26	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	0.69

Densitometry revealed a total of 26 bands within the 2.0–30 kb range (Fig. 2 and Table 1). Patterns obtained from the leukocyte and semen of the same individual as well as both members of the monozygotic twins were indistinguishable from each other. Unrelated individuals were distinguishable from one another based on combinations of detectable bands. Every band of child can be traced to either one of the parents. No mutations were observed. Polymorphism was not limited to males. Three out of the 26 bands, Nos. 8, 23 and 25, appear to be constant in Japanese.

Paternity probability was 0.98 in family 14–16. The ratios of the coincidence probability of individual peaks were between 0.3 and 0.8 in most peaks. The total probability of exclusion was estimated to be $1 - (6.65 \times 10^{-6})$ with the assumption that there is no linkage disequilibrium between bands.

DISCUSSION

A pHY10 fragment contains 64 *TaqI* sites. Thus *TaqI* digestion of pHY10 generates an average of 56 bp fragments, the largest fragment being 221 bp (Nakahori *et al.*, 1986). The major part of the DYZ1 family consists of pHY10 and shows minimal divergence (unpublished data). Thus, it is expected that *TaqI* digestion of human genomic DNA efficiently removes the Y specific DYZ1 family and uncovers normally occult pHY10-related DNA sequences, if any. They can either be Y specific or autosomal in origin. Both satellites II and III are among the sequences related to TTCCA, the main motif of pHY10 (Prosser *et al.*, 1986). According to Fowler *et al.* (1988), "pure" satellite III contains no *TaqI* site although a small proportion of diverged fragments with some sites are included in the satellite III family.

In the present study, *TaqI* digestion revealed minisatellite-like and polymorphic DNA in both males and females when probed with pHY10. Up to 26 bands were resolved, 23 of which showed polymorphism. The overall complexity of the *TaqI*-pHY10 system roughly corresponded to Jeffreys' minisatellite with one of their probes (Jeffreys *et al.*, 1986). Fowler *et al.* (1987, 1988) detected hypervariable polymorphism using satellite III related probes. The probe 228s, among others, detected extensive polymorphic patterns. It is a 332 bp fragment in which TGGAA (TTCCA on the other strand) is the most common pentanucleotide motif and appears 35 times out of possible 66 (53%). The pHY10 includes 229 (32%) TTCCAs out of possible 713 pentanucleotides. They distribute rather evenly in the pHY10.

In the present study, pHY10 was used as a probe in combination with *TaqI* and successfully detected polymorphism of autosomal origin. The probe has been established to be very useful in the sex identification of dried blood stains for forensic purposes (Kobayashi *et al.*, 1988). It is also used in the detection of male cells after bone-marrow transplantation and in the detection of structural rearrangements involving the Y chromosome (Morisaki *et al.*, 1988; Lau, 1985;

Agematsu *et al.*, 1988; Nakahori *et al.*, unpublished data). The probe is easily available either from the ATCC or from the JCRB gene bank, Japan. Now, it can be used to detect either Y or autosomal polymorphism for without constraint of patents. We believe the present technique represents a significant addition to the practical aspects of DNA polymorphism study for forensic, anthropological and human genetic purposes such as zygosity diagnosis, paternity dispute, and other applications of "DNA fingerprinting."

Acknowledgments This study was supported in part by the Chiyoda Mutual Life Foundation, the Ministry of Health and Welfare and the Ministry of Education, Science and Culture of Japan. The authors are indebted to Prof. S.R. Young, Univ. of South Carolina, who has kindly read over the manuscript.

REFERENCES

- Agematsu, K., Koike, K., Morosawa, H., Nakahori, Y., Nakagome, Y. and Akabane, T. 1988. Chondrodysplasia punctata with X:Y translocation. *Human Genet.* **80**: 105-107.
- Bostock, C.J., Gosden, J.R. and Mitchell, A.R. 1978. Localization of a male-specific DNA fragment to a sub-region of the human Y chromosome. *Nature* **272**: 324-328.
- Cooke, H. 1976. Repeated sequence specific to human males. *Nature* **262**: 182-186.
- Fowler, C., Drinkwater, R. and Skinner, J. 1987. Hypervariable lengths of human DNA associated with a human satellite III sequence found in the 3.4 kb Y-specific fragment. *Nucleic Acids Res.* **15**: 3929.
- Fowler, C., Drinkwater, R., Skinner, J. and Burgoyne, L. 1988. Human satellite-III DNA: an example of a "macrosatellite" polymorphism. *Human Genet.* **79**: 265-272.
- Jeffreys, A.J., Wilson, V., Thein, S.L., Weatherall, D.J. and Ponder, B.A.J. 1986. DNA "fingerprints" and segregation analysis of multiple markers in human pedigrees. *Am. J. Hum. Genet.* **39**: 11-24.
- Kobayashi, R., Nakauchi, H., Nakahori, Y., Nakagome, Y. and Matsuzawa, S. 1988. Sex identification in fresh blood and dried bloodstains by a nonisotopic deoxyribonucleic acid (DNA) analyzing technique. *J. Forensic Sci.* **33**: 613-620.
- Lau, Y.F. 1985. Detection of Y-specific repeat sequences in normal and variant human chromosomes using *in situ* hybridization with biotinylated probes. *Cytogenet. Cell Genet.* **39**: 184-187.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. 1982. *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Lab., Cold Spring Harbor, New York.
- Morisaki, H., Morisaki, T., Nakahori, Y., Ogura, H., Kanno, H., Tani, K., Kodo, H., Fujii, H., Asano, S. and Miwa, S. 1988. Genotypic analysis using a Y-chromosome-specific probe following bone-marrow transplantation. *Am. J. Hematol.* **27**: 30-33.
- Nakahori, Y., Mitani, K., Yamada, M. and Nakagome, Y. 1986. A human Y-chromosome specific repeated DNA family (DYZ1) consists of a tandem array of pentanucleotide. *Nucleic Acids Res.* **14**: 7569-7580.
- Prosser, J., Frommer, M., Paul, C. and Vincent, P.C. 1986. Sequence relationships of three human satellite DNAs. *J. Mol. Biol.* **187**: 145-155.