

Two Tetranucleotide Repeats within the Xq21.3/Yp11.2 Human Specific Region of Homology and Their Conservation in Primate Evolution

Heui-Soo Kim¹, Catharine Winstanley², Rekha V. Wadekar¹, Osamu Takenaka³,
Fusako Mitsunaga³, Takashi Kageyama³, Byung-Hwa Hyun⁴,
and Timothy J. Crow^{1*}

¹*Department of Psychiatry, Warneford Hospital, University of Oxford,
Oxford, United Kingdom,*

²*Department of Psychology and Physiology, University of Oxford,
Oxford, United Kingdom,*

³*Department of Cellular and Molecular Biology, Primate Research Institute,
Kyoto University, Inuyama, Japan, and*

⁴*Genetic Resources Center, Korea Research Institute of Bioscience
and Biotechnology, Taejeon, Korea*

ABSTRACT—We locate two tetranucleotide repeat sequences (AT3 and T2C2) between the markers sY44 and sY46 within the Xq21.3/Yp block of homology that has been created since the separation of the chimpanzee and human lineages, and trace their origin in primate evolution. The T2C2 repeat is present only in hominoid primates. The sequence AT3 is present in Old World monkeys but not in New World monkeys, and has been lost in some gibbon species. In the bonobo, the AT3 repeat is the site of a new Alu insertion. These findings and their relationship to the conservation of other markers in this region cast light on the structure of a genomic region that has been subject to change in the course of primate evolution and may include one or more sites of instability.

INTRODUCTION

A general characteristic of mammalian genome organization is the relative stability of the X chromosome across species, particularly when compared to the variation on the Y chromosome. One source of such variation on the Y is X to Y translocation. These translocations may duplicate on the Y genes previously present only on the X. A number of such events are described in the course of mammalian evolution, and have been dated by studies of marker conservation in relation to lineage separations.

The most recent such event in human evolution is the reduplication of a block of sequences in Xq21.3 to Yp11.2, a translocation that was followed by a paracentric inversion of the distal 3.5Mb of the newly translocated block on the Y chromosome (Sargent *et al.*, 1997; Mumm *et al.*, 1997). This translocation separates human from the chimpanzee and is dated at between 3 and 4 million years. The subsequent paracentric inversion (Schwartz *et al.*, 1992) has not been precisely dated but maybe much more recent. These events are interesting

as they have potential relevance to characteristics, including sexual dimorphisms, that have changed in recent primate evolution (Crow, 1993). Some evidence of genetic linkage of psychosis and cerebral asymmetry, (degrees of relative hand skill), a possible correlate of language, has been reported (Laval *et al.*, 1998).

One polymorphic marker (a pentanucleotide repeat) within the Xq21.3/Yp11.2 homologous region has been identified (Chen *et al.*, 1994). Here we report the identification of two tetranucleotide repeats within the Xq21.3 block, assign them a location in relation to the STS content of the recently constructed YAC contig (Sargent *et al.*, 1996), and describe their origin in primate evolution. The findings have relevance to studies of gene expression within the region and to the sequence of evolutionary changes with which the region may be associated.

MATERIALS AND METHODS

YAC analysis and PCR

YAC clones from human Xq21.3, provided by Dr. Nabeel Affara, Cambridge University, were analysed to determine the regional locations of AT3 (ATTT repeat) and T2C2 (TTCC repeat) using the polymerase chain reaction (PCR) method. The PCR primers HS3 (5'-

* Corresponding author: Tel. +44-1865-226474;
FAX. +44-1865-244990.
E-mail: tim.crow@psychiatry.oxford.ac.uk

CAGGGAAGAGGAAAGACTGA-3', bases 115555–115574) and HS4 (5'-CATGTAACCTACCACCAATC-3', bases 115680–115690) from human Xq21.3 for AT3, HS7 (5'-CCCTCAAGAACTAGTCTAGG-3', bases 102346–102366) and HS8 (5'-CTTCTACAGCAGGCCTCATT-3', bases 102551–102570) for T2C2 were designed from the GenBank, Accession No.AC004071. The PCR conditions followed were those of Kim *et al.* (1996).

Primate DNA

DNA was isolated following a standard protocol (Sambrook *et al.*, 1989) from heparinized blood samples from the hominoid primates - bonobo (*Pan paniscus*) - 4 individuals, chimpanzee (*Pan troglodytes*) - 2 individuals, gorilla (*Gorilla gorilla*) - 2 individuals, orangutan (*Pongo pygmaeus*) - 2 individuals, agile gibbon (*Hylobates agilis*) - 5 individuals, siamang (*Hylobates syndactylus*) - 6 individuals, Muller's gibbon (*Hylobates muelleri*) - 4 individuals, Old World monkeys - Japanese monkey (*Macaca fuscata*) - 2 individuals, hamadryas baboon (*Papio hamadryas*) - 2 individuals, African green monkey (*Cercopithecus aethiops*) - 2 individuals, rhesus monkey (*Macaca mulatta*) - 2 individuals, New World monkeys - Squirrel monkey (*Saimiri sciureus*) - 1 individual, night monkey (*Aotus trivirgatus*) - 1 individual, and marmoset (*Callithrix jacchus*) - 1 individual.

Cloning of PCR products

The PCR products were separated on 2% agarose gel, purified using a QIAEX II gel extraction kit (QIAGEN) and cloned into T-khs307 vector (Kim *et al.*, 1998). Plasmid DNA was isolated by the alkali lysis method using the High Pure Plasmid Isolation Kit (Boehringer Mannheim). Individual plasmid DNA was screened for successful insertion by PCR.

DNA sequencing and data analysis

Nucleotide sequences were determined on both strands of plasmid DNA using an automated DNA sequencer (Model 373A, Applied Biosystem) and DyeDeoxy terminator kit with T7 and M13 reverse primers. Sequence analyses were done with the aid of the GCG program (Oxford University).

RESULTS

To determine the regional location of two tetranucleotide-repeat markers, (ATTT)_n and (TTCC)_n, YAC clones from the human Xq21.3 region were subjected to PCR analysis. Using specific primers designed from the GenBank, accession no. AC004071 (170760 bp), two fragments were detected, designated AT3 and T2C2. In an analysis of YACs from the Xq21.3 region (Sargent *et al.*, 1996) we found that PCR product for each of the two tetranucleotides was amplified only from clone 13 (YAC clone 25BE9) that includes the STS markers sY46 and sY47 (Fig. 1). The absence of products from clones 11CH5 (that includes sY44) and 27IG10 (that includes sY46) indicates that AT3 and T2C2 must lie between STSs sY44 and sY46, ie close to marker sY47. PCR analysis of AT3 and T2C2 in various primate species is shown in Fig. 2. PCR products were obtained for all hominoid primates and Old World monkeys except for the agile gibbon and rhesus monkey (147 bp) using the AT3 primers (Fig. 2A), while amplification was only achieved with hominoid primates (226 bp) using T2C2 primers (Fig. 2B).

Interestingly, the PCR product from the DNA of both the male and female bonobo contained 291 bp Alu elements in the AT3 fragments (Fig. 3B). The sequencing data allowed us to confirm the exact site of integration of Alu elements into the bonobo genome. As shown in Fig. 3A, the 12 bp sequence ATTTATTTATTT, which is repeated directly in AT3 fragments, is present only in the bonobo, despite the comprehensive selection of primates examined. This tetranucleotide repeat (ATTT) sequence clearly represents the duplicated target site of integration. To investigate our findings further, DNA from

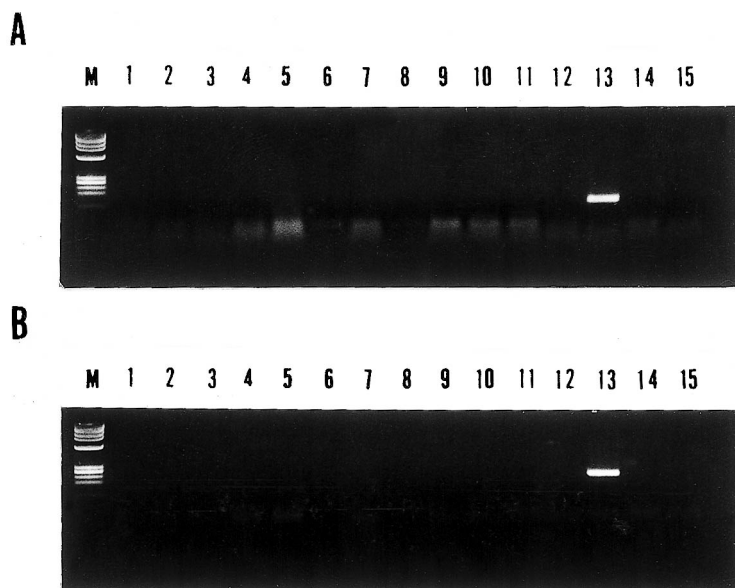


Fig. 1. PCR analysis of YAC clones from the human Xq21.3 region using AT3 (A) and T2C2 primers (B). The lanes are numbered for the YAC clone. 1: 40ID10, 2: 36IA2, 3: 9AD1, 4: 7AA9, 5: 40AD6, 6: 29GA2, 7: 37GC10, 8: 33DC10, 9: 13DE8, 10: 17GB5, 11: 12AH2, 12: 11CH5, 13: 25BE9, 14: 27IG10, 15: 11FC8, and M: marker (ϕ X174/Hae III).

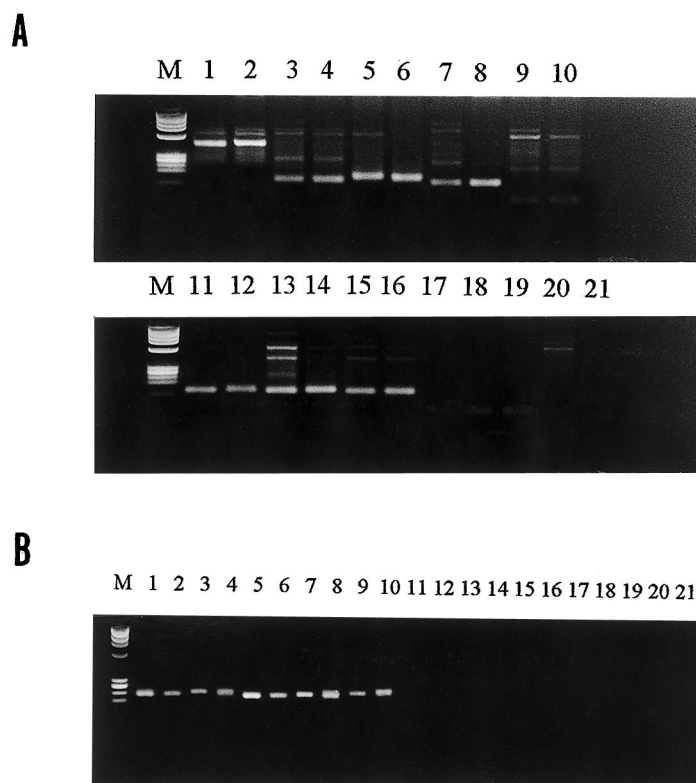


Fig. 2. PCR analysis of genomic DNA in primates using AT3 (A) and T2C2 primers (B). The lanes are numbered for the primate species. 1: bonobo male, 2: bonobo female, 3: chimpanzee male, 4: chimpanzee female, 5: gorilla male, 6: gorilla female, 7: orangutan male, 8: orangutan female, 9: agile gibbon male, 10: agile gibbon female, 11: Japanese monkey male, 12: Japanese monkey female, 13: hamadryas baboon male, 14: hamadryas baboon female, 15: African green monkey male, 16: African green monkey female, 17: rhesus monkey male, 18: rhesus monkey female, 19: Squirrel monkey female, 20: night monkey female, 21: marmoset female, and M: marker (ϕ X174/Hae III).

(A)

BON CCACATAATAACATGACAAAACCTATTAATTAATTAATTTATTTATTTATTTATTTATTT**ATT**TATTT
||||| ||||| | ||| ||| |||||
HUM CCACATAGTAACATGACAAAAC...ATTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT

BON **ATTT** (Alu 291 bp) **ATTTATTTATTT**TAAATTTGTGTGGATTTGATAGAAATGTTCCATACCTT
|||||
HUM ATTT TTAATTTGTGTGGATTTGATAGAAATGTTCCATACCTC

(B)

1	ATTTTTTTTG	AGGCGGAGTC	TCGCTCTGTC	GCCTGGGCTG	GAGTGCAGTG
51	GCGCGATCTC	GGCTCACTGC	AAGCTCCGCC	TCCCGGGTTC	ACGCCATTCT
101	CCTGCCTCAG	CCTCCCGAGT	AGCTGGGACT	ACAGGCGCCC	GCTACCACGC
151	CCGGCTAATT	TTTTGTATTT	TTAGTAGAGA	CGGGGTTTCA	CCGTGTTAGC
201	CAGGATGGTC	TCGATCTCCT	GACCTCGTGA	TCCGCCCCGC	TCGGCCTCCC
251	AAAGTGCTGG	GATTACAGGC	GTGAGCCACC	GCGCCCCGCC	A

Fig. 3. (A) Comparison of nucleotide sequences of the bonobo (BON) and human (HUM) AT3 fragments. The human sequence is taken from GenBank, accession no. AC004071. (B) Nucleotide sequences of 291 bp Alu element of bonobo.

two different bonobo individuals was analysed to determine whether the Alu element was present. The positive results, shown in Fig. 4A (lanes 1 and 2) confirm that the Alu element is apparently present in all bonobo individuals.

As no amplification was achieved with DNA from the agile gibbon (Fig. 2A, lanes 9 and 10), individuals from different subspecies of gibbon (siamang, Muller's gibbon) were investigated. DNA samples were analysed by PCR using the AT3 primers. As a control, the samples were also subjected to analysis using sY46 primers. The results, as shown in Fig. 4, indicate that whereas sY46 amplified well from each of the samples, there was substantial variation using the AT3 primers both within and between subspecies. It seems therefore that the AT3 repeat is present in the siamang and Muller's gibbon but possibly not in the agile gibbon.

In order to determine the evolutionary significance of this variation, other primers designed from neighbouring regions within Xq21.3 were used to analyse the primate DNA (Table 1). The results, as shown in Table 2, indicate that sY24, sY34 and sY43 were clearly amplified in all the primate samples examined, from hominoid species through to New World monkeys. The XY8 primers only produced a positive result in hominoids and Old World monkeys. In addition to T2C2, sY29 and sY48 are also only found in hominoid primates. SY52 is amplified in the bonobo, chimpanzee, gorilla and orangutan

samples, whereas sY47 is present only in the bonobo, chimpanzee and gorilla. Finally, with the exception of the hamadryas baboon, sY46 is present in all the primate samples analysed.

DISCUSSION

Our studies localise the two tetranucleotide repeats to the region between the STS markers sY44 and sY46. This area is within the telomeric portion on the X of the q21.3 region that is represented as a 3.5Mb block in Yp11.2. This region is of interest in man because this block was included in a transposition from the X to the Y between the separation of the lines that led to the chimpanzee and human, and was involved in a subsequent paracentric inversion (Sargent *et al.*, 1996; Mumm *et al.*, 1997). Hence, the origin of the X-Y homology is recent but the precise organization and earlier evolutionary history of the sY44 to sY46- region (that includes sY47) raises some problems. The order of sY46 and sY47, as reported by Sargent *et al.* (1996), is the same on the X and Y chromosomes, but was reported as reversed by Mumm *et al.* (1997) who raised the question of genomic instability in this region. Furthermore sY47 is located on several YACs (eg 161C3, 9BH10 and 13FB8) that also include the marker GMGXY12 that according to the comparative studies of

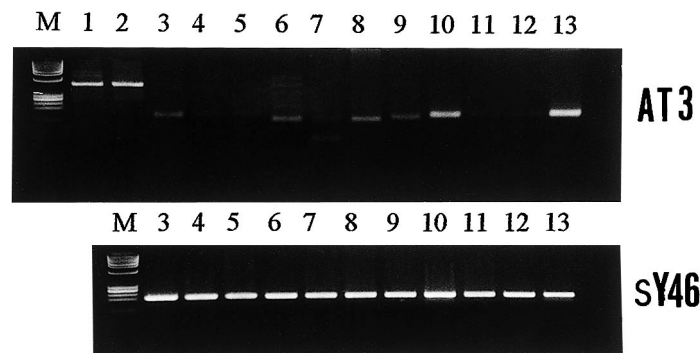


Fig. 4. PCR analysis of genomic DNA in the bonobo and gibbon subspecies using AT3 and sY46 primers. The lanes are numbered for the primate species. 1: bonobo male A; 2: bonobo male B; 3: agile gibbon male A, 4: agile gibbon female B, 5: agile gibbon male C, 6: siamang male A, 7: siamang female B, 8: siamang male C, 9: siamang female D, 10: Muller's gibbon male A, 11: Muller's gibbon female B, 12: Muller's gibbon male C, 13: Muller's gibbon female D, and M: marker (ϕ X174/Hae III).

Table 1. STS and primer sequence used for PCR analysis

STS	Primer A	Primer B	Ann. temp.(°C)	Size
XY8	5'-AGAGATCCATAATCTTTTGC-3'	5'-CCCAGAGACCAGAACATTTT-3'	56	200
sY24	5'-TATGATGTTAGCCATGGGCT-3'	5'-ATGAGGCCAGCATTACTCTG-3'	57	373
sY29	5'-CCCTATCAGCTCTGTAGCCA-3'	5'-AGCAGCACAGAACCCCTTATG-3'	57	283
sY34	5'-ACTGTGAGCGAGCTGAAAAT-3'	5'-TGCAGCCTTGTGAACCAATTA-3'	58	293
sY43	5'-CCTTTTCATGACGTATAGCTGG-3'	5'-AAAGTTTTGAATACCAAGAGAAGG-3'	56	300
sY47	5'-ATCTGTGAAGCACCCACTGT-3'	5'-CCTGTTGCAAATGACAAATG-3'	58	174
sY46	5'-TATTTGTGGGAAAATCTGGC-3'	5'-CCAGATGCCACAGTTCACTA-3'	52	202
sY48	5'-GAGGCAATCACAAATTCTGG-3'	5'-TGGCTGAATGATAACCCATT-3'	52	329
sY52	5'-CTCAAAATTTCCGTGTGCTT-3'	5'-CGCCAAAACAGGAGATTAGA-3'	52	292
AT3	5'-CAGGGAAGAGGAAAGACTGA-3'	5'-CATGTAACCTACCACCACAATC-3'	53	147
T2C2	5'-CCCTCAAGAACTAGTCTAGG-3'	5'-CTTCTACAGCAGGCCTCATT-3'	55	226

Table 2. PCR analysis using human primers in primates

	BON		CHI		GOR		ORA		AGI		JMO		HBA		GMO		RMO		SMO	NMO	MAR
	m	f	m	f	m	f	m	f	m	f	m	f	m	f	m	f	m	f	f	f	f
XY8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	X	X	X
sY24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
sY29	0	0	0	0	0	0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X
sY34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
sY43	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
sY47	0	0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
sY46	0	0	0	0	0	0	0	0	0	0	0	0	X	X	0	0	0	0	0	0	0
sY48	0	0	0	0	0	0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X
sY52	0	0	0	0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X
AT3	0*	0*	0	0	0	0	0	0	X	X	0	0	0	0	0	0	X	X	X	X	X
T2C2	0	0	0	0	0	0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X

* Alu element was inserted in AT3 of the bonobo.

Species names are abbreviated as follows. BON, bonobo; CHI, chimpanzee; GOR, gorilla; ORA, orangutan; AGI, agile gibbon; JMO, Japanese monkey; HBA, hamadryas baboon; GMO, African green monkey; RMO, rhesus monkey; SMO, squirrel monkey; NMO, night monkey; MAR, marmoset. m, male; f, female; 0, amplification; X, no amplification.

Lambson *et al.* (1992) and YAC and sequence analyses of Sargent *et al.* (1996) must have colonised the Y chromosome at least twice in primate evolution. Yet further evidence of a possible rearrangement was noted by Mumm *et al.* (1997) in relation to sY50, the marker adjacent to GMGXY12 on the X chromosome, that was originally reported (Foote *et al.*, 1992) between sY49 and sY51 on the Y chromosome but was mapped between sY43 and sY44 on the X chromosome by Sargent *et al.* (1996) and Mumm *et al.* (1997).

Thus there is evidence of instability in this region between and perhaps within species. In this context we find that the evolutionary origins of these two tetranucleotide repeats is different. AT3 is present in hominoid primates (with the exception of some gibbon species) and Old World monkeys (Japanese monkey, hamadryas baboon, African green monkey) with the exception of the rhesus monkey, but absent in New World monkeys. These findings are compatible with an origin after the separation of the New and Old World monkeys with subsequent loss in some gibbon species and in the rhesus monkey. By contrast to AT3, T2C2 is present only in hominoid primates. Within this super-family it is found to be highly conserved across species and sex.

In the bonobo (but not in *Pan troglodytes*) the AT3 repeat sequence has been the target for the insertion of a 291 bp Alu element (Fig. 3). This observation adds to evidence for continued insertions of Alu elements in the course of the primate radiation (Arcot *et al.*, 1995) and reinforces the concept that such insertions are associated with specific target sequences, including AT rich repeats (Daniels and Deininger, 1985; Zuliani and Hobbs, 1990).

Across primate species, we find a high degree of conservation of some parts of the Xq21.3/Yp homologous blocks (Table 2). Thus the regions identified by the markers sY24, sY34, and sY43 are present in each of the groups we have studied. SY46 was present in each of the primate species with the exception of the hamadryas baboon, in whose lineage there may have been an unspecified evolutionary event.

By contrast, and in agreement with Lambson *et al.* (1992), we find that XY8 (a marker for the proximal region of homology on the Y chromosome) is present only in hominoids and Old World monkeys. This indicates that this region developed after the speciation of Old World from New World monkeys.

A set of markers – sY29, sY48, and T2C2 – is present only in the hominoids primates. Because they are separated by markers (eg sY34 and 43, and sY46) that were present much earlier in evolution these markers may represent three separate additions to the X chromosome before, or in the course of the hominoid radiation. This result implies that these regions were created after hominoid primates evolved from Old World monkeys. An alternative explanation is that the primers, being designed from human DNA, failed to amplify in Old World and New World monkeys as a result of sequence variation that distinguishes hominoids from other primates.

In summary we report two tetranucleotide repeats within the Xq21.3/Yp11.2 region of homology and have outlined their separate origins in primate evolution. The transposition and the paracentric inversion of the Xq21.3 homologous block on the Y chromosome may have implications for characteristics that are human specific. Thus the origin and structure of potential sites of instability and change within this region are of interest in relation to the course of primate including human evolution.

ACKNOWLEDGMENTS

This work was supported by grants from the UK Medical Research Council and the Stanley Foundation. We thank N A Affara for a gift of YAC clones and R M Ridley, H F Baker and N Mundy for samples of primate DNA.

REFERENCES

- Arcot SS, Wang GR, Weber JL, Deininger PL, Batzer MA (1995) Alu repeats: a source for the genesis of primate microsatellites. *Genomics* 29: 136–144

- Chen HM, Lowther W, Avramopoulos D, Antonorakis SE (1994) Homologous loci DXYS156X and DXYS156Y contain a polymorphic pentanucleotide repeat (TAAAA) and map to human X chromosome and Y chromosome. *Hum Mutation* 4: 208–211
- Crow TJ (1993) Sexual selection, Machiavellian intelligence and the origins of psychosis. *Lancet* 342: 594–598
- Daniels GR, Deininger PL (1985) Integration site preferences of the Alu family and similar repetitive DNA sequences. *Nucleic Acids Res* 13: 8939–8954
- Foots S, Vollrath D, Hilton A, Page DC (1992) The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* 258: 60–66
- Kim H-S, Chen Y, Lonai P (1998) Complex regulation of multiple cytohesin like genes in murine tissues and cells. *FEBS Lett* 433: 312–316
- Kim H-S, Hirai H, Takenaka O (1996) Molecular features of the TSPY gene of gibbons and Old World monkeys. *Chrom Res* 4: 500–506
- Lambson B, Affara NA, Mitchell M, Ferguson-Smith MA (1992) Evolution of DNA sequence homologies between the sex chromosomes in primate species. *Genomics* 14: 1032–1040
- Laval SH, Dann J, Butler RJ, Loftus J, Rue J, Leask SJ, Bass N, Comazzi M, Vita A, Nanko S, Shaw S, Peterson P, Shields G, Smith AB, Stewart J, DeLisi LE, and Crow TJ (1998) Evidence for linkage to psychosis and cerebral asymmetry (relative hand skill) on the X chromosome. *Am. J. Med. Genet. (Neuropsychiatric Genet)* (in press)
- Mumm S, Molini B, Terrell J, Srivastava A, Schlessinger D (1997) Evolutionary features of the 4Mb Xq21.3 XY homology region revealed by a map at 60-kb resolution. *Genome Res* 7: 307–314
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
- Sargent CA, Briggs H, Chalmers IJ, Lambson B, Walker E, Affara NA (1996) The sequence organization of Yp/proximal Xq homologous regions of the human sex chromosomes is highly conserved. *Genomics* 32: 200–209
- Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Distèche C, McGillivray B, De la Chapelle A, Page DC (1998) Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE-LINE recombination. *Hum Mol Genet* 7: 1–11
- Zuliani G, Hobbs HH (1990) A high frequency of length polymorphisms in repeated sequences adjacent to Alu sequences. *Am J Hum Genet* 46: 963–969

(Received November 24, 1998 / Accepted January 15, 1999)