

Centromere Emergence in Evolution

Mario Ventura, Nicoletta Archidiacono, and Mariano Rocchi¹

Sezione di Genetica-DAPEG, 70126 Bari, Italy

Evolutionary centromere repositioning is a paradox we have recently discovered while studying the conservation of the phylogenetic chromosome IX in primates. Two explanations were proposed: a conservative hypothesis assuming sequential pericentric inversions, and a more challenging assumption involving centromere emergence during evolution. The complex evolutionary history showed by chromosome IX did not allow us to clearly distinguish between these two hypotheses. Here we report comparative studies of chromosome X in two lemur species: the black lemur and the ringtailed lemur. The X chromosome is telocentric in the black lemur and almost metacentric in the ringtailed lemur. The marker order along these chromosomes, however, was found to be perfectly colinear with humans. Our data unequivocally point to centromere emergence as the most likely explanation of centromere repositioning.

Human centromeric and pericentromeric regions have been shown to be highly plastic (Archidiacono et al. 1995; Eichler et al. 1999; Jackson et al. 1999; Horvath et al. 2000), in sharp contrast to the relative stability of the rest of the genome (Kaessmann et al. 1999). We have recently reported an additional puzzling feature of centromeric regions, the centromere repositioning paradox, discovered while studying the cytogenetic conservation of phylogenetic chromosome IX in primates (Montefalcone et al. 1999). In some species the centromere shows an evolutionary history independent from the surrounding markers. Its position can be reconciled with these markers only by assuming additional, peculiar pericentric inversions (which we named flip-flop inversions) or by hypothesizing evolutionary emergence of centromeres. Distinction between these two alternatives is critical to our understanding of the processes of chromosomal evolution. The flip-flop hypothesis does not imply any new biological mechanism and would just add further support to the opinion that intrachromosomal rearrangements are very frequent in evolution (Muller et al. 2000). Evolutionary appearance of centromeres, on the contrary, would represent a novel, unpredicted biological property.

The X chromosome is considered the most evolutionary conserved chromosome among mammals (Ohno 1973; Chowdhary et al. 1998). Also, its submetacentric shape, as it appears in humans, is highly conserved and very likely represents the mammalian ancestral form. This assumption is strongly supported by the perfect conservation of marker order, including the centromere, between cat (*Felix catus*) and human X chromosomes (Murphy et al. 1999). We reasoned that the striking X chromosome evolutive conservation was

a unique opportunity that would facilitate the testing of centromere repositioning mechanisms. Comparative studies on marker order conservation among the X chromosome of humans (*Homo sapiens*, HSA) and two Lemuridae species: *Eulemur macaco* (EMA; black lemur) and *Lemur catta* (LCA; ringtailed lemur) were performed. These two primates were selected for the study because the morphology of the X chromosome in these two species is quite different from humans. Despite these differences, no marker order discrepancy was observed. Our results strongly indicate, therefore, that the repositioning took place via centromere emergence.

RESULTS

Chromosome X appears telocentric in EMA and almost metacentric in LCA (Fig. 1a). To investigate their evolutionary conservation, we performed fluorescence in situ hybridization (FISH) analysis using different kinds of probes. Human Partial Chromosome Paints (PCP) specific for Xp (#364) and Xp11.2→11.3 (#377) were initially used (Fig. 1b,c). Despite the striking difference in centromere location, no pericentric inversion was detected. This conclusion was also supported by the use of PCPs #102 (Xq; data not shown) and #118 (Xq24→28; Fig. 1d). To investigate marker order conservation in more detail, a panel of 19 human probes spanning the chromosome X from the tip of Xp to the tip of Xq was then used (Table 1). A probe containing the *PGPL* gene, located inside the PseudoAutosomal Region 1 (PAR1) and mapping 80 to 110 kb from Xpter (Gianfrancesco et al. 1998), was used as a telomeric Xp marker. A cosmid clone (U130F6) containing the *HSPRY3* gene was used as a marker of the Xqter region. *HSPRY3* maps inside the PAR2, approximately 250 kb from the Xq telomere (Ciccodicola et al. 2000). All the probes were found to hybridize in EMA and LCA as if their X chromosomes were perfectly colinear with the human X. Examples of these experiments are reported

¹Corresponding author.

E-MAIL rocchi@biologia.uniba.it; FAX 39-080-544-3386.

Article and publication are at www.genome.org/cgi/doi/10.1101/gr.152101.

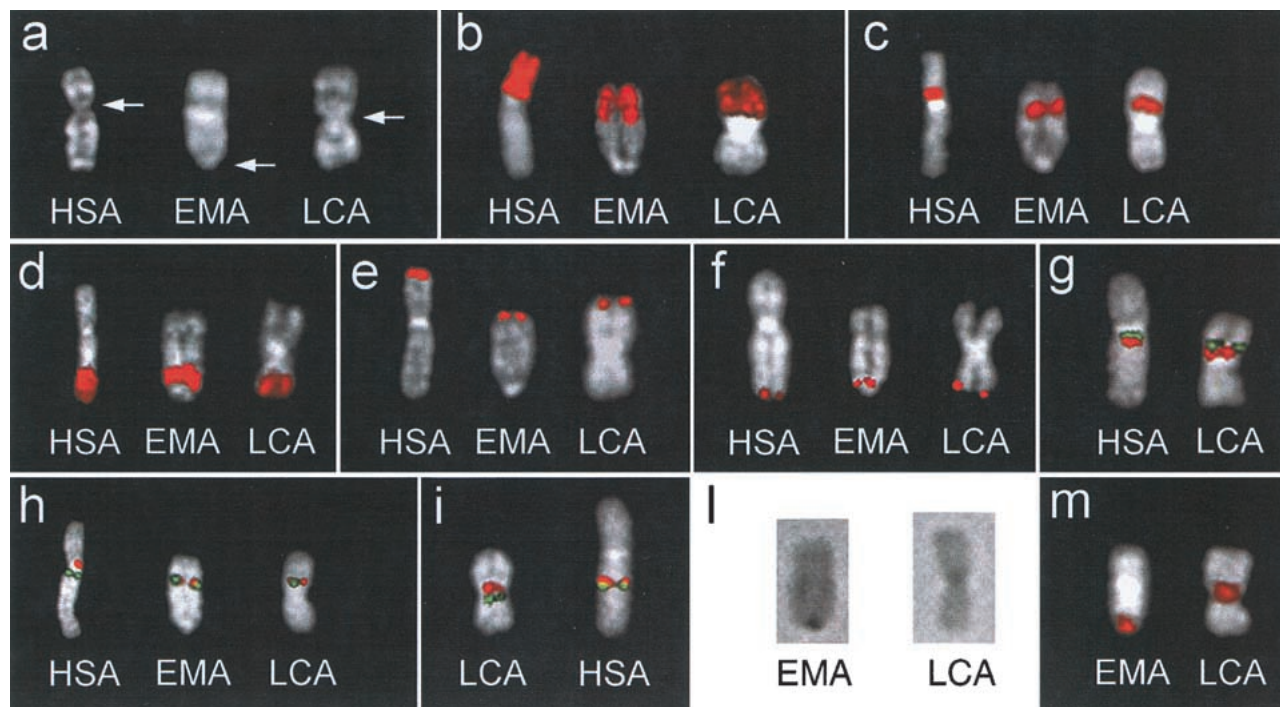


Figure 1 (a) Q-banded chromosome X from *Homo sapiens* (HSA), *Eulemur macaco* (EMA; black lemur), and *Lemur catta* (LCA, ringtailed lemur). The arrows indicate the centromere. EMA and LCA chromosomes are oriented, in all images, upside-down to match the orientation of human X. (b–f) Examples of fluorescence in situ hybridization (FISH) experiments using PCP #364 (Xp; b), PCP #377 (Xp11.2–Xp11.3; c), PCP #118 (Xq24–28; d), PCP #PGPL probe (PAR1; e), and U130F6 cosmid clone (*HSPRY3* gene, PAR2; f). (g) Cohybridization experiment using probes bA235K20 (green signal) and dJ1015G2 (red signal). Their orientation is identical in both species. (h) Probes dJ598A24 (red) and bA235K20 (green) flank the human centromere. In EMA and LCA, they appear to be almost completely overlapping. (i) Probes dJ1015G2 (red) and dJ715D6 (green) surround the LCA centromere and are well separated by the centromeric heterochromatin. In HSA, they appear much closer. (j) The C-banded chromosome X of EMA and LCA. (m) The signal of DOP-PCR amplified products from microdissected material of the EMA (left) and LCA (right) centromeric regions hybridized to EMA and LCA metaphases, respectively. Only chromosome X is shown (see text for details).

in Figure 1e and 1f. The centromere position in EMA was found to be telomeric to all the studied markers, including the *HSPRY3* gene (Fig. 1f).

The centromere location in LCA was further investigated using cohybridization experiments, because its different position in respect to HSA could be easily explained by a pericentric inversion of a relatively small chromosomal region. The experiments reported in Figure 1g are crucial in this respect. BAC bA235K20 maps at Xq11.2, below the human centromere; PAC dJ1015G2B, at Xq13 in HSA, is located above the LCA centromere. Both probes are within the hypothetical inverted region; therefore, they would necessarily detect the inversion. Their FISH signals clearly show that their orientation is maintained in both HSA and LCA, thus proving that the difference in centromere position was not caused by a pericentric inversion.

Additional human PAC or BAC probes close to the HSA and LCA centromeres were identified by querying the Sanger Centre database. FISH analysis indicated that dJ598A24 and bA235K20 were very close to the human centromere on the Xp and Xq side, respectively (Fig. 1h). They were cohybridized on EMA and LCA

metaphases. The two signals were found to nearly overlap in both species (Fig. 1h). An additional experiment was performed using PACs dJ1015G24 and dJ715D6. These two probes are located at the opposite sides of the LCA centromere, well separated by the centromeric heterochromatin (Fig. 1i, left). In HSA, they appear much closer (Fig. 1i, right).

C-banding technique was used to investigate the nature of the cytogenetic material located at the centromere of EMA and LCA chromosome X. The results are shown in Figure 1j. The two centromeres appear to contain a heterochromatic block comparable in size to the average of the autosomal chromosomes. To test the possibility that this region was the result of a transposition from the ancestral centromeric area or from an autosomal region, we microdissected both centromeres. About 20 samples for each species were collected, DOP-PCR amplified, biotin-labeled, and hybridized on metaphase preparations of EMA, LCA, and HSA. Labeled DOP-PCR material from EMA gave a strong signal on all EMA chromosomes, with the exception of chromosome 1. The signal on chromosome X is shown in Figure 1m. No signal was detected on

Table 1. Probes Used in the Study

	Probe	STS	cM	cR	Mapping
COS	<i>PGPL</i>	80-110 kb from Xpter			Xp22.3
YAC	749G10	DXS1043	17.7	7.65	Xp22.2
YAC	681F6	DXS987	25.5		Xp22.1
YAC	789C8	DXS989	40.6		Xp22.1
YAC	965C6	DXS997	47.0	25.91	Xp21
YAC	649F6	DXS991	86.9		Xp11.2
PAC	dj598A24	DXS6967			Xp11
HSA centromere					
BAC	bA235K20	WI-14304	90.8	169.61	Xq11.2
PAC	dj1015G24	DXS8109	97.9		Xq13
YAC	933D12	DXS6809		245.07	Xq21
YAC	748C4	DXS8077	108.1		Xq21
		DXS6740		245.17	
LCA centromere					
PAC	dj715D6				Xq21.3
YAC	963G9	DXS1072	122.8		Xq23
		AFMA275ZC9		266.47	
YAC	923D3	DXS1001	139.4	275.97	Xq24
YAC	719H8	DXS984	159.5		Xq26/2q32
		WI-6425		299.53	
YAC	960C9	DXS1205	163.7		Xq26
		DXS6798		302.92	
YAC	878A5	DXS1215	183.8	312.58	Xq27
YAC	849E12	DXS8103	192.5		Xq27-28/13q14
COS	<i>HSPRY3</i>	250kb from Xqter			Xq28 (PAR2)
EMA centromere					

Most of the probes are YACs of the CEPH megalibrary, obtained from the YAC Screening Center (Milan). They were selected after a careful evaluation of their position on the genetic (cM) and radiation hybrid (cR) map (MIT database).

HSA and LCA chromosomes. The labeled DOP-PCR products from LCA centromere lit up the centromere of the X chromosome of LCA (Fig. 1m) and the DAPI-positive interstitial heterochromatic blocks present on some LCA chromosomes (data not shown). In EMA, a clear FISH signal was detected on the pericentromeric heterochromatic block of chromosomes 6 and 3 (data not shown). No signal was found in HSA. An hypothetical transposition of autosomal material into the X chromosome of LCA was also searched for by hybridizing a complete panel of human WCPs on LCA metaphases. Painting signal of LCA chromosome X was obtained only by the WCP derived from human chromosome X (data not shown). Reciprocal chromosome painting analysis of EMA and HSA chromosomes has been performed by Muller et al. (1997). They report a perfect equivalence of the euchromatic content of human and EMA X chromosomes.

Remains of alphoid sequences at the ancestral X centromere of EMA and LCA were searched for by using p82H α satellite plasmid and a pool of human alphoid clones as probes, hybridized at low stringency. The pool included the alphoid probe specific for chromosome X. Probe p82H recognizes, at low stringency, all the human centromeres (Alexandre et al. 1987). Both experiments failed to detect any signal on EMA and LCA chromosomes.

DISCUSSION

We have studied the evolutionary conservation of chromosome X in HSA and two Lemuridae species, EMA and LCA, using a panel of appropriate FISH probes. The chromosome X is telocentric in EMA and almost metacentric in LCA. Despite the very apparent morphological differences, our results indicate that marker order is perfectly conserved. Our data, therefore, are an additional example of centromere repositioning in primates. Thanks to the striking evolutionary conservation of chromosome X in mammals, the present study represented a unique opportunity of testing the different hypotheses that could be formulated to explain the centromere repositioning phenomenon. The perfect marker conservation in both EMA and LCA does not fit with the flip-flop hypothesis. Indeed, no intermediate form (caused by a single inversion) of the X chromosome has been documented in primates.

Centromere repositioning not mediated by chromosome rearrangement can be hypothesized to occur either through a transposition event that would insert centromeric sequences into an euchromatic region or through neocentromere emergence. FISH experiments using DOP-amplified products of the microdissected centromeric regions of EMA and LCA X chromosomes were used to substantiate a hypothetical insertional transposition from a different centromere. No hints in

favor of a transposition event including euchromatic DNA were found. This conclusion is also supported by the systematic use of human chromosome paints on LCA chromosomes (present data) and by reverse painting data reported by Muller et al. (1997) on LCA. A transposition event restricted to heterochromatic material, however, could not be discarded with certainty in EMA. Transposition of centromeric sequences into a distinct centromere has been documented in a prenatal diagnosis case (Verlinsky et al. 1995). As far as we know, however, no examples are available, in mammals, of neocentromere seeding caused by insertional transposition of a centromeric block into a noncentromeric region. On the contrary, since the first convincing report of neocentromere occurrence in humans (du Sart et al. 1997; Barry et al. 2000), several additional cases have been described (Warburton et al. 2000). In one instance, the neocentromere showed a normal familial segregation (Tyler-Smith et al. 1999). Neocentromeres also have been occasionally documented in *Drosophila* (Williams et al. 1998). Neocentromere emergence, therefore, seems to be the most likely explanation of the centromere repositioning we have documented.

An intriguing issue arising from this conclusion is the loss or gain of heterochromatic material shown, respectively, by the ancestral centromere locus and by the newly formed centromeres. The interpretation of the data is that the heterochromatic material, usually associated with higher eukaryote centromeric regions, has disappeared in both EMA and LCA from the site where the ancestral centromere was located, as a consequence of the deep remodeling of the region following centromere repositioning. Indeed, C-banded material was not detected at the ancestral centromere loci, whereas it was evident at the active centromere of both EMA and LCA. We can hypothesize that heterochromatic material gradually accumulates at the newly formed centromere following neocentromere seeding. Evolution of neocentromeres toward complexity is probably driven by a specific selective pressure (Eichler 1999). Unfortunately, the intermediate evolutionary events of such heterochromatization processes are not tractable. A clear example, however, of the degradation process affecting an inactivated centromere is provided by the ancestral centromere at 2q21 in HSA, which became inactive following the telomere-telomere fusion of the two ancestral phylogenetic chromosomes 11p and 11q (Ijdo et al. 1991). Despite its recent origin (at most 3–5 million years ago), relics of alphoid sequences are barely detectable at 2q21 (Avarello et al. 1992; Baldini et al. 1993), and there is no evidence of C-banded material.

Studies supporting the view that centromere repositioning is a relatively widespread phenomenon in mammalian evolution are starting to emerge (Monte-

falcone et al. 1999; Band et al. 2000; Iannuzzi et al. 2000; Muller et al. 2000). Altogether, these data indicate that centromere repositioning played a role in mammalian genome evolution.

METHODS

Metaphase spreads were obtained from PHA-stimulated peripheral blood lymphocytes from a normal human donor and from fibroblast cell lines of EMA and LCA. PCPs were generated in our laboratory (see our Web site, http://www.biologia.uniba.it/rmc/1_hy-tutti/1-2_hy_diagrams/X.html). WCPs, derived from flow-sorted chromosomes, were a gift from Dr. N. Carter (Sanger Centre). WCP specific for chromosome 22 was obtained by *Alu*-PCR amplification of DNA from a somatic cell hybrid retaining chromosome 22 as the only human chromosome (see our Web site). YAC probes of CEPH megalibrary were obtained from the YAC Screening Centre (Milan). PAC and BAC probes were from the Sanger Centre (<http://webace.sanger.ac.uk>). The "bA" (RPCI BAC library) or "dJ" (RPCI PAC library) prefix, used by the Sanger Centre to identify the source library, was maintained. Probes *PGPL* and *HSPRY3* were a generous gift of A. Ciccodicola (IIGB, Naples). FISH experiments have been described in detail elsewhere (Montefalcone et al. 1999). Microdissection and DOP-PCR amplification procedures were performed according to Meltzer et al. (1992).

ACKNOWLEDGMENTS

The financial support of cofin2000-MURST, AIRC, and Telethon (Grant E.672 to M.R. and E.962 to N.A.) is gratefully acknowledged.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Archidiacono, N., Antonacci, R., Marzella, R., Finelli, P., Lonoce, A., and Rocchi, M. 1995. Comparative mapping of human alphoid sequences in great apes using fluorescence in situ hybridization. *Genomics* **25**: 477–484.
- Alexandre, C., Miller, D.A., Mitchell, A.R., Warburton, D.A., Gersen, S.L., Disteche, C., and Miller, O.J. 1987. p82H identifies sequences at every human centromere. *Hum. Genet.* **77**: 46–50.
- Avarello, R., Pedicini, A., Caiulo, A., Zuffardi, O., and Fraccaro, M. 1992. Evidence for an ancestral alphoid domain on the long arm of human chromosome-2. *Hum. Genet.* **89**: 247–249.
- Baldini, A., Ried, T., Shridhar, V., Ogura, K., D'Aiuto, L., Rocchi, M., and Ward, D.C. 1993. An alphoid DNA sequence conserved in all human and great ape chromosomes: Evidence for ancient centromeric sequences at human chromosomal regions 2q21 and 9q13. *Hum. Genet.* **90**: 577–583.
- Band, M.R., Larson, J.H., Rebeiz, M., Green, C.A., Heyen, D.W., Donovan, J., Windish, R., Steining, C., Mahyuddin, P., Womack, J.E., et al. 2000. An ordered comparative map of the cattle and human genomes. *Genome Res.* **10**: 1359–1368.
- Barry, A.E., Bateman, M., Howman, E.V., Cancilla, M.R., Tainton, K.M., Irvine, D.V., Saffery, R., and Choo, K.H. 2000. The 10q25 Neocentromere and its inactive progenitor have identical primary nucleotide sequence: Further evidence for epigenetic modification. *Genome Res.* **10**: 832–838.
- Ciccodicola, A., D'Esposito, M., Esposito, T., Gianfrancesco, F., Migliaccio, C., Miano, M.G., Matarazzo, M.R., Vacca, M., Franz, A., Cuccurese, M., et al. 2000. Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal

- region. *Hum. Mol. Genet.* **9**: 395–401.
- Chowdhary, B.P., Raudsepp, T., Fronicke, L., and Scherthan, H. 1998. Emerging patterns of comparative genome organization in some mammalian species as revealed by Zoo-FISH. *Genome Res.* **8**: 577–589.
- du Sart, D., Cancilla, M.R., Earle, E., Mao, J., Saffery, R., Tainton, K.M., Kalitsis, P., Martin, J., Barry, A.E., and Choo, K.H.A. 1997. A functional neocentromere formed through activation of a latent human centromere and consisting of non- α -satellite DNA. *Nat. Genet.* **16**: 144–153.
- Eichler, E.E. 1999. Repetitive conundrums of centromere structure and function. *Hum. Mol. Genet.* **8**: 151–155.
- Eichler, E.E., Archidiacono, N., and Rocchi, M. 1999. CAGGG repeats and the pericentromeric duplication of the hominoid genome. *Genome Res.* **9**: 1048–1058.
- Gianfrancesco, F., Esposito, T., Montanini, L., Ciccodicola, A., Mumm, S., Mazzarella, R., Rao, E., Giglio, S., Rappold, G., and Forabosco, A. 1998. A novel pseudoautosomal gene encoding a putative GTP-binding protein resides in the vicinity of the Xp/Yp telomere. *Hum. Mol. Genet.* **7**: 407–414.
- Horvath, J.E., Viggiano, L., Loftus, B.J., Adams, M.D., Archidiacono, N., Rocchi, M., and Eichler, E.E. 2000. Molecular structure and evolution of an α satellite/non- α satellite junction at 16p11. *Hum. Mol. Genet.* **9**: 113–123.
- Iannuzzi, L., Di Meo, G.P., Perucatti, A., Incarnato, D., Schibler, L., and Cribiu, E.P. 2000. Comparative FISH mapping of bovid X chromosomes reveals homologies and divergences between the subfamilies Bovinae and Caprinae. *Cytogenet. Cell Genet.* **89**: 171–176.
- Ijdo, J.W., Baldini, A., Ward, D.C., Reeders, S.T., and Wells, R.A. 1991. Origin of human chromosome 2: An ancestral telomere-telomere fusion. *Proc. Natl. Acad. Sci.* **88**: 9051–9055.
- Jackson, M.S., Rocchi, M., Thompson, G., Hearn, T., Crosier, M., Guy, J., Kirk, D., Mulligan, L., Ricco, A., Piccininni, S., et al. 1999. Sequences flanking the centromere of human chromosome 10 are a complex patchwork of arm-specific sequences, stable duplications and unstable sequences with homologies to telomeric and other centromeric locations. *Hum. Mol. Genet.* **8**: 205–215.
- Kaessmann, H., Heissig, F., von Haeseler, A., and Paabo, S. 1999. DNA sequence variation in a non-coding region of low recombination on the human X chromosome. *Nat. Genet.* **22**: 78–81.
- Meltzer, P.S., Guan, X.Y., Burgess, A., and Trent, J.M. 1992. Rapid generation of region specific probes by chromosome microdissection and their application. *Nat. Genet.* **1**: 24–28.
- Montefalcone, G., Tempesta, S., Rocchi, M., and Archidiacono, N. 1999. Centromere repositioning. *Genome Res.* **9**: 1184–1188.
- Muller, S., O'Brien, P.C.M., Ferguson-Smith, M.A., and Wienberg, J. 1997. Reciprocal chromosome painting between human prosimians (*Eulemur macaco macaco* and *E.fulvus mayottensis*). *Cytogenet. Cell Genet.* **78**: 260–271.
- Muller, S., Stanyon, R., Finelli, P., Archidiacono, N., and Wienberg, J. 2000. Molecular cytogenetic dissection of human chromosomes 3 and 21 evolution. *Proc. Natl. Acad. Sci.* **97**: 206–211.
- Murphy, W.J., Sun, S., Chen, Z.Q., Pecon-Slattery, J., and O'Brien, S.J. 1999. Extensive conservation of sex chromosome organization between cat and human revealed by parallel radiation hybrid mapping. *Genome Res.* **9**: 1223–1230.
- Ohno, S. 1973. Ancient linkage groups and frozen accidents. *Nature* **244**: 259–262.
- Tyler-Smith, C., Gimelli, G., Giglio, S., Florida, G., Pandya, A., Terzoli, G., Warburton, P.E., Earnshaw, W.C., and Zuffardi, O. 1999. Transmission of a fully functional human neocentromere through three generations. *Am. J. Hum. Genet.* **64**: 1440–1444.
- Verlinsky, Y., Ginsberg, N., Chmura, M., Freidline, M., White, M., Strom, C., and Kuliev, A. 1995. Cross-hybridization of the chromosome 13/21 α satellite DNA probe to chromosome 22 in the prenatal screening of common chromosomal aneuploidies by FISH. *Prenat. Diagn.* **15**: 831–834.
- Warburton, P.E., Dolled, M., Mahmood, R., Alonso, A., Li, S., Naritomi, K., Tohma, T., Nagai, T., Hasegawa, T., Ohashi, H., et al. 2000. Molecular cytogenetic analysis of eight inversion duplications of human chromosome 13q that each contain a neocentromere. *Am. J. Hum. Genet.* **66**: 1794–1806.
- Williams, B.C., Murphy, T.D., Goldberg, M.L., and Karpen, G.H. 1998. Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat. Genet.* **18**: 30–37.

Received June 12, 2000; accepted in revised form January 17, 2001.