**Running title:**

**Chromosome-specific DNA Repeat Probes**

# Chromosome-specific DNA Repeat Probes

Adolf Baumgartner [1,2], Jingly Fung Weier [1,2] and Heinz-Ulrich G. Weier [2, ✉]

[1] Dept. of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, CA, USA, and

[2] Life Sciences Division, University of California, E.O. Lawrence Berkeley Natl. Lab., Berkeley, CA, USA.

✉Correspondence to

H.-U. Weier, Department of Genome Biology, Life Sciences Division, MS 74-157

University of California, O.E. Lawrence Berkeley National Laboratory

1 Cyclotron Road, Berkeley, CA 94720, USA.

Phone (510) 486 5347, fax (510) 486 5343, e-mail: ulliweier@hotmail.com

Prepared for: Journal of Histochemistry and Cytochemistry, Full Report

**Summary**

Fluorescence in situ hybridization (FISH) has gained increasing popularity as a highly sensitive technique to study cytogenetic changes in the research environment as well as in the clinical laboratory. Today, hundreds of commercially available probes serve the basic needs of the biomedical research community. Widespread applications, however, are often limited by the lack of appropriately labeled, specific nucleic acid probes. We describe two rapid approaches to the preparation of chromosome-specific probe DNAs and readily available methods to label the probes with the reporter molecules of choice. Notably, the techniques allow preparation of highly specific DNA repeat probes suitable for enumeration of chromosomes in interphase cell nuclei or tissue sections without a need for chromosome enrichment by flow cytometry and sorting or molecular cloning. Examples of production of DNA repeat probes specific for either human chromosome 17 or 18 also demonstrate that the entire process from probe concept and design to successful hybridization can be completed in the laboratory in just a few days without a need for highly specialized equipment.

**KEY WORDS:** Cytogenetics, chromosome analysis, numerical aberrations, chromosome enumeration, DNA repeats, DNA probes, fluorescence in situ hybridization (FISH)

## Introduction

In many fields such as cytogenetics, cancer research, forensics, environmental cytotoxicology or preimplantation genetic diagnosis (PGD), fluorescence *in situ* hybridization (FISH) replaced the previously used techniques, i.e., Giemsa-banding analyses, to evaluate numerical and structural aberrations (Gray et al. 1991; Weier et al. 1991). Paralleling an increasing demand for DNA probes, availability, costs, specificity and efficiency of these probes have become important parameters. Taking advantage of resources generated in the course of the International Human Genome Project, yeast artificial chromosome (YAC) and bacterial artificial chromosome (BAC) libraries (Weissenbach et al. 1992; Ioannou et al. 1994) have been used for the generation of chromosome- or locus-specific DNA probes (Fung et al. 1998, 2001; Weier et al. 1991, 1994). The present report describes the targeted amplification of known chromosome-specific repeat sequences and the utilization of BAC clones that map in close proximity to a centromere as a template for a PCR-based generation of chromosome-specific repeat probes, which yield bright centromeric FISH signals on the targeted chromosomes.

Tandemly repeated alphoid DNA in the centromeres region of human chromosomes consists of distinct subfamilies in a number exceeding the total number of chromosomes (Jorgensen 1997). Alpha satellites most often contain monomer variants that differ from the consensus sequence by up to 40% (Rosandic et al. 2003). At least 33 different alphoid subfamilies have been identified because of their organization into different tandem arrays. Some of these subfamilies are specific for a single chromosome, whereas others are common to a small group of

chromosomes (Rosandic et al. 2003). The basic repeat units of alpha satellite DNA are divergent AT-rich monomers of approximately 171-bp organized throughout the genome in chromosome-specific higher-order-repeat units (Manuelidis 1978a, 1978b; Mitchell et al. 1985; Willard 1991). There are five alpha satellite suprachromosomal families displaying characteristic types of monomers, all descending from two ancestral types, A and B (Alexandrov et al. 1988, 2001). Selection-driven evolution formed the distinction between both prototypes. Every non-acrocentric human chromosome possesses at least one chromosome-specific family of alpha satellite, which is defined by a unique higher-order repeated unit originating from simple ancestral repeats, one to five monomers long. These have spread through the genome by cycles of inter-chromosomal transfers and amplifications. Each suprachromosomal family is characterized by a unique set of monomeric types, which often alternate regularly within genomic sequences (Romanova et al. 1996). The primate X centromeres, for instance, appears to have evolved through repeated proximal expansion events by unequal recombination occurring within the central, active region of the centromeric DNA (Schueler et al. 2005).

Evolutionary dispersion of alphoid repeats from their centromeric locus to further distal loci enables the use of proximal BACs as PCR templates to generate centromere-specific DNA probes. Human pericentromeric regions, i.e. regions proximal to the centromere, are hot spots for duplication events as well as prone to genetic instability (Eichler 1998). Considerable variability of segmental duplications

within 5 Mbp from the centromere exists in this region of certain chromosomes occasionally leaving clusters of alphoid sequences proximal to the centromere (She et al. 2004). Using this evolutionary variability to our advantage, we present a versatile and rapid way to generate DNA repeat probes by PCR amplification of DNA from BACs located proximal to the centromere of the target chromosome.

Most BAC clones have been mapped to the euchromatic portion of the human genome. The BAC clones, especially those in close proximity of the centromeres, frequently contain single copies of DNA repeats such as satellite DNA being derived from evolutionary expansion of centromeric regions during the phylogenetic progression of the centromeres. With an increasing number of fully sequenced BAC clones and, very importantly, their DNA sequences publicly available, simple database searches allow rapid identification of promising clones to generate chromosome-specific DNA repeat probes. By using BAC clones from pericentromeric regions of the p- and q-arms as a template for PCR amplification with alpha-satellite specific primers, complexity of the different subfamilies of alpha-satellite can be reduced which can result in chromosome-specific centromeric DNA probes as shown below. Without this reduction of complexity cross-hybridization leads to labeling of other chromosomes with similar alphoid repeats rendering the probe unsuitable for detecting a single chromosome pair. For cytogenetic diagnostics, however, the rapid production of bright and specific probes for fluorescence in situ hybridization is of utmost importance to reliably detect numerical and structural chromosomal abnormalities.

## Material and Methods

*PCR and DNA labeling*

For PCR, 0.1 µg genomic DNA or BAC DNA were used as amplification templates. The PCR reactions (50 µl) were performed using 0.02 U/µl Taq Polymerase (Invitrogen) in 1x PCR buffer (Invitrogen), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, and 0.6 µM of the forward and reverse primers (Qiagen, Alameda, CA). As primers, P17H8-F1 and P17H8-R1-r for chromosome 17 (P17H8-F1: TGAACATTCCTATTGATAGAGCAG, P1H8-R1-r: CTCCAGTTTTTATGTGAC-CATAA, product size: 799 bp) (Waye and Willard 1986) as well as pYAM9-60F1 and pYAM9-60R2-r for chromosome 18 (pYAM9-60F1: CTGCAGCGTTCTGAGAAA-CATC, pYAM9-60R2-r: GCGGGAATTCATACAAATTGCAG, predicted product size: 1311 bp) (Alexandrov et al. 1991) were employed. For the chromosome 18-specific PCR, JumpStart™ Taq polymerase (Sigma, St. Louis, MO) was used instead of the normal Taq polymerase (Roche Diagnostics Corp., Indianapolis, IN) for a more efficient hot start PCR. After an initial denaturation step of 1 min at 95 °C, 35 cycles followed consisting of 30 sec at 95 °C (1 min ramp to 54 °C), 1 min at 54 °C (1 min ramp to 72 °C) and 3 min at 72 °C (30 sec ramp to 95 °C). Finally, after 10 min at 72 °C, the reactions were kept at room temperature for a few hours. The PCR products were confirmed on a 2% agarose gel (Fig. 1a) by applying 5 µl of the PCR reaction mixed with 1 µl 6x sucrose (4g sucrose in 10 ml water).

Random priming was employed to directly label the PCR-derived probe DNA with fluorochrome-conjugated nucleotides. For indirect labeling, biotin-14-dCTP or

digoxigenin-11-dUTP (Roche Diagnostics) were incorporated into the DNA using a commercial kit (BioPrime Kit, Invitrogen).

*Selection of BAC clones*

The BAC clones 285M22 (GenBank accession number AC131274) and 18L18 (GenBank accession number AC136363) from library RP11 (Invitrogen, Gaithersburg, MD) were chosen based on information available at UC Santa Cruz (http://genome.ucsc.edu/cgi-bin/hgGateway) (Fig. 1e) and the National Center for Biotechnology and Information (NCBI) (http://www.ncbi.nlm.nih.gov/mapview/map_search. cgi?taxid=9606). We selected BAC clones reported to contain alphoid sequences, extracted and labeled the DNA, before PCR and FISH were performed.

*Preparation of DNA probes*

BAC DNA was isolated using a standard alkaline lysis DNA extraction protocol (Birnboim and Doly 1979, Weier et al. 1994, 1995a). The isolation was done from 10 ml bacterial cultures in Luria-Bertani (LB) medium containing 12.5 µg/ml chloramphenicol. Briefly, the cells were treated with lysozyme and lysed under alkaline conditions. The DNA was then precipitated in isopropanol, washed once in 70% cold ethanol, and resuspended in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). Finally, the DNA was extracted with phenol and chloroform (pH 7; Pierce, Rockford, IL), precipitated with 2.5 volumes of ethanol, and resuspended in double distilled sterile water (approximately 20-40 µl depending on the size of the pellet)(Weier et al. 1995b).

*Preparation of metaphase spreads*

Metaphase spreads were prepared from phytohemagglutinin (PHA)-stimulated peripheral blood lymphocytes from a karyotypically normal male (Fung et al. 2001). White blood cells were cultured for 76 hours in RPMI 1640 medium. The cells were then synchronized by treatment with $10^{-7}$ M methotrexate (MTX) for 15 hours to inhibit DNA replication, followed by $10^{-5}$ M thymidine for 5 hours to release the cells synchronously from the MTX-induced block. Colcemid (1 mg/ml) was added during the final 30 minutes of thymidine release. Lymphocytes were fixed in methanol and acetic acid (3:1), and dropped on pre-cleaned microscope slides. The slides were allowed to air-dry in a Thermotron environmental chamber (Thermotron Industries, Holland, MI) at 25 °C and 47.5% humidity (Fung et al. 2001).

*FISH*

For FISH, 1 μl of either probe, 1 μl of human Cot-1 DNA® (1 mg/ml, Invitrogen), 1 μl of salmon sperm DNA (10 mg/ml; Invitrogen), and 7 μl of the hybridization master mix [78.6% formamide, 14.3% dextran sulfate in 1.43x SSC, pH 7.0 (20x SSC is 3 M sodium chloride, 300 mM tri-sodium citrate)] were thoroughly mixed and subsequently denatured at 76 °C for 10 min. Then, the hybridization mixture was incubated at 37 °C for 30 min allowing the Cot-1 DNA® to pre-anneal with the probes. In parallel, the slides were denatured for 3 min at 76 °C in 70% formamide/2x SSC, pH 7.0, dehydrated in 70%, 85%, and 100% ethanol for 2 min each step, and allowed to air-dry. Then, the hybridization mix was carefully applied

onto the slides, covered with 22x22 $mm^2$ cover-slips and sealed with rubber cement. Slides were incubated overnight in a moist chamber. After carefully removing the rubber cement, the slides were immersed in 2x SSC at room temperature until the cover-slips slid off. Subsequently, the slides were washed in 0.1x SSC at 43 °C for 2 min, then incubated in PNM [5% nonfat dry milk (Carnation), 1% sodium azide, in PN buffer (0.1 M sodium phosphate buffer, pH 8.0, 1% Nonidet-P40)] for 10 min at room temperature. Bound probes were detected with fluorescein-conjugated avidin (Vector labs, Burlingame, CA) and rhodamine-labeled anti-digoxigenin antibodies (Roche Diagnostics). Finally, the slides were mounted in 4',6-diamidino-2-phenylindole (DAPI, 0.5 µg/ml; Calbiochem, La Jolla, CA) in antifade solution (Weier et al. 1995a,b).

*Image acquisition and analysis*

Fluorescence microscopy was performed on a Zeiss Axioskop microscope equipped with a filter set for simultaneous observation of Texas Red/rhodamine and FITC, and a separate filter for DAPI detection (ChromaTechnology, Brattleboro, VT). Images were collected using a cooled CCD camera (CCD-1300DS, VDS Vosskuehler, Osnabrück, Germany). Further processing and printing of the images were done using image processing software Adobe Photoshop (Adobe Systems Inc., San Jose, CA).

## Results

In Fig 1a, the PCR products can be seen on the gel. For the chromosome 18 PCR (Fig. 1a, lane L1), multiple bands were observed with the strongest between 603 bp and 750 bp, approximately with 700bp. Additionally, bands of PCR amplified DNA with a higher molecular weight are visible at around 900 bp and 1300 bp. There is also a noticeable band with roughly 350 bp. The high-molecular smear at the top of the lane could be a result of overloading the gel. For chromosome 18, we can confirm that genomic DNA is sufficient for PCR amplification with alpha-satellite specific primers to create a chromosome-specific bright FISH probe (Fig. 1b). Nice, bright signals are visible on chromosome 18.

After PCR amplification and labeling, both selected pericentromeric BACs containing chromosome 17 specific alpha-satellite sequences yield in bright signals at the centromere of chromosome 17. The probe derived from a PCR using BAC RP11-285M22 as a template (Fig 1d) was slightly brighter than that using BAC RP11-18L18 (result not shown). When employing genomic DNA instead of the BAC DNA in an equivalent PCR reaction, multiple cross-hybridizing signals on various chromosomes can be observed (Fig. 1c) within the metaphase spread. On the gel, the difference between the two lanes L2 and L3 (Fig. 1a) is evidently the additional band at around 350 bp, which results in massive cross-hybridization when using genomic DNA instead of the pericentromeric BAC as a template. Using either BAC RP11-285M22 (Fig 1a, lane L3) or RP11-18L18 (Fig 1a, lane L4) results in only one PCR product at 799 bp.

Figure 1e shows the results of our databank search for RP11-285M22 (Genbank accession number AC131274) highlighted in red. The shorter BAC RP11-18L18 (Genbank accession number AC136363) can be seen just below. Both BAC are located in the pericentromeric area of 17p11.1-11.2. This region contains two monomeric alpha satellites clusters M2 and M3 (Rudd et al. 2006), which are the targets of our chromosome 17 alpha satellite-specific PCR primers.

## Discussion

Fluorescence *in situ* hybridization is a widely accepted molecular cytogenetic technique which greatly relies on chromosome- or gene-specific DNA probes to produce accurate signals when bound to their targets. Despite a variety of commercially available FISH probes, multicolor FISH assays create an increasing demand for specifically designed probes for scientific and clinical research; especially probes for disease loci or those labeled with suitable fluorochrome labels. Here, we propose a rapid, inexpensive and reproducible procedure to generate large quantities of chromosome-specific probe DNAs able to results in specific, bright hybridization signals. Two approaches for the rapid PCR synthesis of probes were used: a.) generating the probe DNA with target-specific oligonucleotides from genomic template DNA in cases when a chromosome-specific repeat sequence is known, and b.) using a BAC DNA as template to reduce complexity when less specific PCR primers are employed for in vitro DNA amplification.

In a targeted approach, we searched public databases to find primers specific for the chromosome 18 derived alphoid sequence (Alexandrov et al. 1991): a 22-mer forward primer (named pYAM9-60F1) and a 23-mer reverse primer (pYAM9-60R2-r). While we used high molecular weight genomic DNA as a PCR template to generate the specific chromosome 18 probe DNA, it is very likely that other DNA templates such as cells from cultures, biopsies or hair follicles lead to similar results. PCR products generated with the primer set pYAM9-60F1/R2-r appeared on agarose gels as a single band in the target site region plus some amount of high molecular weight

DNA fragments (Fig. 1a, lane L1). The resulting DNA probe labeled with FITC showed a unique chromosome-specific hybridization signal at the centromere of chromosome 18 (Fig. 1a). The higher order alphoid centromeric repeats of chromosome 18 apparently are unique to this chromosome and not to produce any cross-hybridization when used in FISH.

The first draft of mapping the human genome does not cover the centromeric regions of all chromosomes or the heterochromatic regions proximal to the centromeres of chromosomes 1, 9, 16, and 19. This explains why there are hardly BACs available for the centromere regions. However, DNA repeats cluster of the centromere region can be found dispersed proximal to, but outside the centromere within the pericentromeric regions of the chromosome. These chromosome-specific alphoid subsets are generally characterized by a higher order repeat, which is composed of diverged monomers in tandemly repeated long arrays (Rocchi et al. 1991).

Pericentromeric regions are defined as sequences extending from the centromere to the first cytogenetic band on a chromosome arm (Eichler et al. 1998). In contrast to the highly conserved repetitive sequences within the human centromeres (Grady et al. 1992), pericentromeric sequences are enriched for inter- and/or intrachromosomally duplicated DNA (37.3%), single copies (49.7%), and satellite sequences (13%). Chromosome 18, for instance, consists of large tracts of duplications within its pericentromeric regions (Mudge and Jackson 2005). These duplications are due to the abiding evolution of repeated DNA sequences by

unequal crossing-over (Schueler et al. 2005; Smith 1976; Willard 1991). Hence, with the number of fully sequenced BAC clones progressively increasing, the BAC clones mapping to either side of the centromere may contain interspersed alphoid repeats within small clusters (see Fig. 1e) and can serve as PCR templates to amplify large amounts of chromosome-specific probe DNA.

For chromosome 17, using genomic DNA is not sufficient to generate a chromosome-specific DNA probe (Fig. 1c). The primers, P17H8-F1 and P17H8-R1, used seemed to be only specific for a complete family of alphoid repeats. Five suprachromosomal families of alpha-satellite DNA can be found on human chromosomes originating from different evolutionary events. While chromosome 18 shows two distinct families 2 as well as 4 and 5, chromosome 17 expresses the suprafamilies 3 and 4 (Alexandrov et al. 2001) with two distinct classes of monomeric alpha-satellite in the centromeric region, at which the M3 monomeric alpha-satellite on band 17p11 is more closely related to higher order alpha-satellite (Rudd et al. 2006) (Fig. 2). Thus, instead of using genomic DNA, BACs proximal to the centromere of chromosome 17 with single copies of a subfamily of alphoid DNA repeats (Rudd et al. 2006) have been employed. Searching public databases allows rapid identification of BAC sequences of suitable clones (Fig. 1e), e.g. the public databases of the University of California Santa Cruz and NCBI, National Center for Biotechnology Information of the National Institute of Health. Two pericentromeric BACs have been mapped to 17p11.1-11.2, RP11-285M22 at the position 22.032-22.197 Mbp and RP11-18L18 at position 22.098-22.170 Mbp, respectively. The

latter co-aligns with RP11-285M22, but is shorter in length. In a PCR with the P17H8-F1 and P1H8-R1-r primers and the selected BACs as templates, DNA probes were generated specifically for the centromere of chromosome 17 (Fig. 1d). Preferably, BAC clones can be directly used for the PCR amplification without isolating the BAC DNA after being grown in LB medium.

Different human chromosomes 17 in the population are characterized by distinct alpha satellite haplotypes, distinguished by the presence of variant repeat forms that have precise monomeric deletions. These repeat units evolve principally along haplotype lineages (Warburton and Willard 1995). Therefore, by decreasing the complexity of these repeat units in the PCR amplification of alphoid sequences the specificity of the generated chromosome-specific DNA probes for the targeted chromosome will increase. Conversely, the variation in these repeats will inevitably generate PCR amplified DNA probes, which demonstrate cross-hybridization to alpha-satellites of other chromosomes. The same strategy of reducing complexity was previously achieved by selecting a subclone containing predominantly monomer 1 or the higher repeat order of the chromosome 17 alphoid repeat (Meyne and Moyzis 1989).

In summary, the rapid approach described here provides large amounts of DNA for labeling with the reporter molecule of choice by either one of the readily available techniques, i.e., random priming, nick translation, tailing or amination. The resulting chromosome-specific DNA probes have been successfully used for FISH yielding bright unambiguous signals in interphase and metaphase cells.

**Acknowledgements**

This work was supported in part by NIH grants CA80792 and CA88258, and a grant from the Director, Office of Energy Research, Office of Health and Environmental Research, U. S. Department of Energy, under contract DE-AC-03-76SF00098. Additional support was provided by grants from the United States Army Medical Research and Material Command, United States, Department of the Army (DAMD17-99-1-9250, DAMD17-00-1-0085). A.B. was supported in part by a grant from the UC Discovery Program. J.F. was supported in part by NIH grants HD045736 and HD041425.

## DISCLAIMER

This publication was made possible by funds received from the Cancer Research Fund, under Interagency Agreement #97-12013 (University of California, Davis, contract #98-00924V) with the Department of Health Services, Cancer Research Section. Mention of trade name, proprietary product or specific equipment does not constitute a guaranty or warranty by the Department of Health Services, nor does it imply approval to the exclusion of other products. The views expressed herein represent those of the authors and do not necessarily represent the position of the State of California, Department of Health Services.

**Literature Cited**

Alexandrov IA, Mitkevich SP, Yurov YB (1988) The phylogeny of human chromosome specific alpha satellites. Chromosoma 96:443-453

Alexandrov IA, Mashkova TD, Akopian TA, Medvedev LI, Kisselev LL, Mitkevich SP, Yurov YB (1991) Chromosome-specific alpha satellites: two distinct families on human chromosome 18. Genomics 11:15-23

Alexandrov I, Kazakov A, Tumeneva I, Shepelev V, Yurov Y (2001) Alpha-satellite DNA of primates: old and new families. Chromosoma 110:253-266

Birnboim HC, Doly J (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res 7:1513-1523

Eichler EE (1998) Masquerading repeats: paralogous pitfalls of the human genome. Genome Res 8:758-762

Eichler EE, Hoffman SM, Adamson AA, Gordon LA, McCready P, Lamerdin JE, Mohrenweiser HW (1998) Complex beta-satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12. Genome Res 8:791-808

Fung J, Hyun W, Dandekar P, Pedersen RA, Weier HUG (1998) Spectral Imaging in Preconception/Preimplantation Genetic Diagnosis (PGD) of Aneuploidy: Multi-Colour, Multi-Chromosome Screening of Single Cells. J Ass Reprod Genet 15:322-329

Fung J, Weier H-UG, Pedersen RA, Zitzelsberger HF (2001) Spectral Imaging Analysis of Metaphase and Interphase Cells. In FISH Technology-Springer Lab Manual (B Rautenstrauss and T Liehr, Eds). Springer Verlag, Berlin. Chapter 29, pp. 363-387

Grady DL, Ratliff RL, Robinson DL, McCanlies EC, Meyne J, Moyzis RK (1992) Highly conserved repetitive DNA sequences are present at human centromeres. Proc Natl Acad Sci U S A 89:1695-1699

Gray JW, Lucas J, Kallioniemi O, Kallioniemi A, Kuo WL, Straume T, Tkachuk D, Tenjin T, Weier HU, Pinkel D (1991) Applications of fluorescence in situ hybridization in biological dosimetry and detection of disease-specific chromosome aberrations. Prog Clin Biol Res 372:399-411

Ioannou P, Amemiya C, Garnes J, Kroisel P, Shizuya H, Chen C, Batzer M, De Jong P (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nat Genet 6:84-89

Jorgensen AL (1997) Alphoid repetitive DNA in human chromosomes. Dan Med Bull 44:522-534

Manuelidis L (1978a) Chromosomal localization of complex and simple repeated human DNAs. Chromosoma 66:23-32

Manuelidis L (1978b) Complex and simple sequences in human repeated DNAs. Chromosoma 66:1-21

Meyne J, Moyzis RK (1989) Human chromosome-specific repetitive DNA probes: targeting in situ hybridization to chromosome 17 with a 42-base-pair alphoid DNA oligomer. Genomics 4:472-478

Mitchell AR, Gosden JR, Miller DA (1985) A cloned sequence, p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. Chromosoma 92:369-377

Mudge JM, Jackson MS (2005) Evolutionary implications of pericentromeric gene

expression in humans. Cytogenet Genome Res 108:47-57

Rocchi M, Archidiacono N, Ward DC, Baldini A (1991) A human chromosome 9-specific alphoid DNA repeat spatially resolvable from satellite 3 DNA by fluorescent in situ hybridization. Genomics 9:517-523

Romanova LY, Deriagin GV, Mashkova TD, Tumeneva IG, Mushegian AR, Kisselev LL, Alexandrov IA (1996) Evidence for selection in evolution of alpha satellite DNA: the central role of CENP-B/pJ alpha binding region. J Mol Biol 261:334-340

Rosandic M, Paar V, Gluncic M, Basar I, Pavin N (2003) Key-string algorithm--novel approach to computational analysis of repetitive sequences in human centromeric DNA. Croat Med J 44:386-406

Rudd MK, Wray GA, Willard HF (2006) The evolutionary dynamics of {alpha}-satellite. Genome Res 16:88-96

Schueler MG, Dunn JM, Bird CP, Ross MT, Viggiano L, Rocchi M, Willard HF, Green ED (2005) Progressive proximal expansion of the primate X chromosome centromere. Proc Natl Acad Sci U S A 102:10563-10568

She X, Horvath JE, Jiang Z, Liu G, Furey TS, Christ L, Clark R, Graves T, Gulden CL, Alkan C, Bailey JA, Sahinalp C, Rocchi M, Haussler D, Wilson RK, Miller W, Schwartz S, Eichler EE (2004) The structure and evolution of centromeric transition regions within the human genome. Nature 430:857-864

Smith GP (1976) Evolution of repeated DNA sequences by unequal crossover. Science 191:528-535

Warburton PE, Willard HF (1995) Interhomologue sequence variation of alpha satellite DNA from human chromosome 17: evidence for concerted evolution

along haplotypic lineages. J Mol Evol 41:1006-1015

Waye JS, Willard HF (1986) Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. Mol Cell Biol 6:3156-3165

Weier HU, Kleine H-D, Gray JW (1991) Labeling of the centromeric region on human chromosome 8 by in situ hybridization. Hum Genet 87:489-494

Weier HU, Rosette CD, Matsuta M, Zitzelsberger H, Matsuta M, Gray J (1994) Generation of highly specific DNA hybridization probes for chromosome enumeration in human interphase cell nuclei: isolation and enzymatic synthesis of alpha satellite DNA probes for chromosome 10 by primer directed DNA amplification. Meth Mol Cell Biol 4:231-248

Weier HUG, Rhein AP, Shadravan F, Collins C, Polikoff D (1995a) Rapid physical mapping of the human trk proto-oncogene (NTRK1) gene to human chromosome 1q21-22 by P1 clone selection, fluorescence in situ hybridization (FISH) and computer-assisted microscopy. Genomics 26, 390-393

Weier HU, Wang M, Mullikin JC, Zhu Y, Cheng JF, Greulich KM, Bensimon A, Gray JW (1995b) Quantitative DNA fiber mapping. Hum Mol Genet 4:1903-1910

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, Lathrop M (1992) A second-generation linkage map of the human genome. Nature 359:794-801

Willard HF (1991) Evolution of alpha satellite. Curr Opin Genet Dev 1:509-514

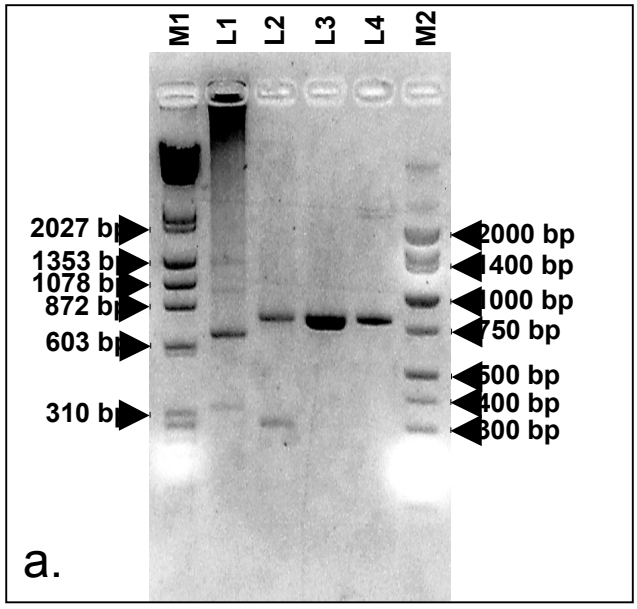
## Figure Legends

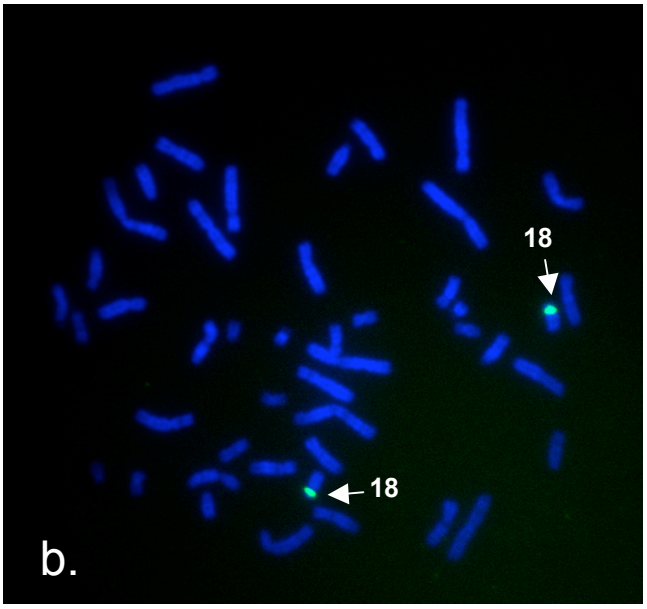
### Figure 1:

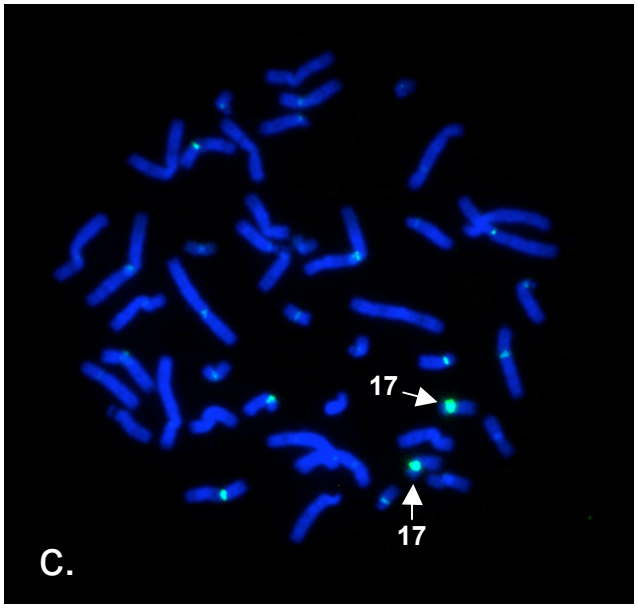
**a.** Agarose gel showing the PCR products, lane L1: chromosome 18 PCR, genomic DNA as template; L2: chromosome 17 PCR, genomic DNA as template; L3: chromosome 17 PCR, BAC RP11-258M22 as template; and L4: chromosome 17 PCR, BAC RP11-18L18 as template. Lanes M1 and M2 are showing two markers, a combined λ Hind III / ΦX174 Hae III marker and a Lo$^{TM}$ marker (Bionexus, Oakland, CA), respectively. **b.** After FISH, bright specific signals on chromosome 18 centromere can be observed using the probe PCR-generated with genomic DNA as a template. **c.** Using genomic DNA to generate a specific probe for chromosome 17 results in more than 2 signals cross-hybridizing to other chromosome centromeres, which renders this probe useless for FISH. **d.** Increased specificity of the chromosome 17 – specific probe when using BAC RP11-385M22 as a PCR template. Bright and very specific signals on both homologues can be seen. **e.** Golden Path of the UCSC genome database showing BACs RP11-285M22 (AC131274, highlighted by the red bar) and RP11-18L18 (AC136363, orange bar) and their relative position on chromosome 17 at p11.1-11.2 as well as the repeat elements within this pericentromeric region. The monomeric alpha-satellites are clearly within the BAC RP11-285M22 sequence.

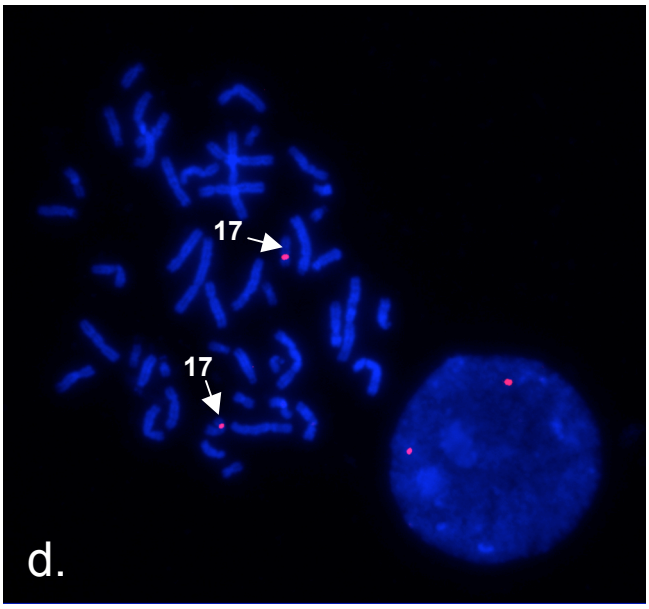
**Figure 2:**

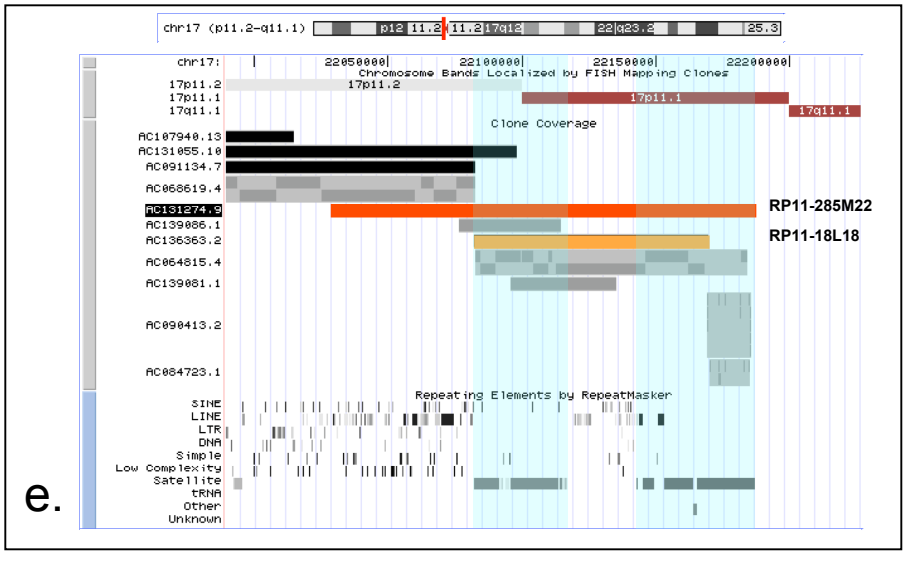
Schematic representation of alphoid repeats (vertical bars) and monomeric repeat clusters (boxes) in the pericentromeric region of human chromosome 17.
A.) Ideogram. B.) Distribution of alphoid satellite repeats. See text for description of regions a-c. Please note that the centromeric heterochromatin in region A is not drawn to scale. Region a: centromere (D17Z1 and D17Z1-B) consisting of tandemly repeated alpha satellites, higher-order repeats; region b: pericentromeric region, clusters of monomeric (M1-M4) and single repeat alpha satellites can be found; region c: this region consists of single copy repeats and genes, isolated single alpha satellite repeats can be found.

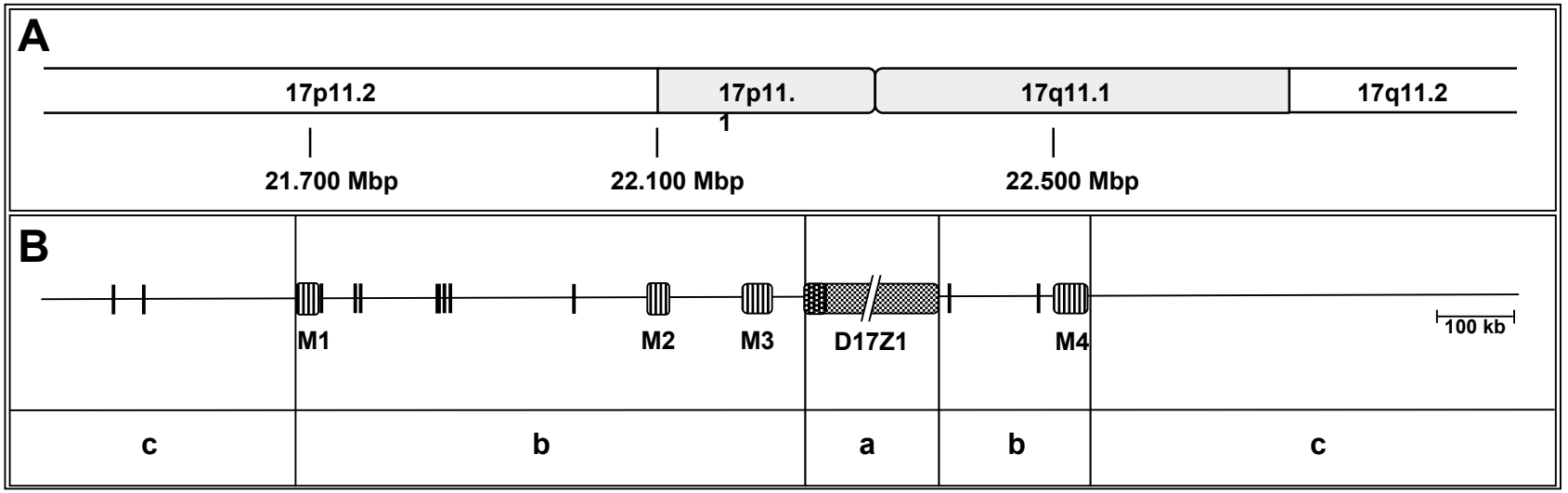

M1
L1
L2
L3
L4
M2
2027 bp
1353 bp
1078 bp
872 bp
603 bp
310 bp
2000 bp
1400 bp
1000 bp
750 bp
500 bp
400 bp
300 bp
a.


18
18
b.


17
17
c.


17
17
d.


chr17 (p11.2-q11.1)
chr17:
22050000
22100000
22150000
22200000
Chromosome Bands Localized by FISH Mapping Clones
17p11.2
17p11.1
17q11.1
Clone Coverage
AC107940.13
AC131055.10
AC091134.7
AC068619.4
AC131274.9
AC139086.1
AC136363.2
AC064815.4
AC139081.1
AC090413.2
AC084723.1
RP11-285M22
RP11-18L18
Repeating Elements by RepeatMasker
SINE
LINE
LTR
DNA
Simple
Low Complexity
Satellite
tRNA
Other
Unknown
e.


A
17p11.2
17p11.1
17q11.1
17q11.2
21.700 Mbp
22.100 Mbp
22.500 Mbp
B
M1
M2
M3
D17Z1
M4
100 kb
c
b
a
b
c