Newly identified repeat sequences, derived from human chromosome 21qter, are also localized in the subtelomeric region of particular chromosomes and 2q13, and are conserved in the chimpanzee genome

H.-S. Park a, M. Nogami b, K. Okumura b, M. Hattori a, Y. Sakaki a, c, A. Fujiyama a, d, e

a RIKEN Genomic Sciences Center, c/o Kitasato University, 1-15-1 Kitasato, Sagamihara, Kanagawa 223-8555, Japan
b Mie University, Faculty of Bioreources, 1515 Kamihama, Tsu, Mie 514-8507, Japan
c National Institute of Genetics, 1111 Yata, Mishima, Shizuoka 411-8540, Japan

d Present address: Department of Genetic Information, Division of Molecular Life Science, Tokai University School of Medicine, Isehara, Kanagawa 259-1195, Japan

e E-mail: afujiyam@lab.nig.ac.jp

Received 19 April 2000
Edited by Horst Feldmann

Abstract Subtelomeric regions have been a target of structural and functional studies of human chromosomes. Markers having a defined structure are especially useful to such studies. Here, we report 93 bp tandem repeat sequences found in the subtelomeric region of human chromosome 21q. They were also detected in the telomeric region of several other chromosomes. Interestingly, the repeat was also found in the 2q13 region which is known to be a position of chromosomal fusion, a major difference between the human and chimpanzee karyotypes. To the best of our knowledge, this repetitive sequence is a new member of human subtelomeric interspersed repeats. © 2000 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Human chromosome; Telomere; Repeat sequence; Primate; Subtelomeric interspersed repeat

1. Introduction

The chromosomal region adjacent to the telomeric repeat array is commonly referred to as the subtelomeric region, and is known to contain a variety of repeat sequences called telomere-associated sequences, most of which are species-specific or shared only by closely related species [1–4]. In human cells, variable repeat sequences have also been found in the region, and these are collectively referred to as subtelomeric interspersed repeats (STIR) [5]. For example, families of 29 bp, 37 bp, and 61 bp repeats present in TelBam 3.4 and TelSat2.0 sequences that are distributed among several chromosomes [6]. It has been reported that the repeating structure at the ends of chromosomes might function in inter-chromosomal rearrangements during meiosis and mitosis, and might also act as a buffer zone to protect genes near the telomere from truncation caused by low telomerase activity [7,8].

In this study, we report a new STIR sequence composed of a tandemly repeated 93 bp sequence covering a region of about 1 kb, mapped to a location 7.5 kb from the telomeric array of the long arm of human chromosome 21. Fluorescence in situ hybridization (FISH) analysis revealed a specific distribution of the repeat units among chromosomes. To the best of our knowledge, this 93 bp repetitive sequence is a new human STIR. In addition, the repeat was also found in the genomic DNA of other species including chimpanzee.

2. Materials and methods

Construction of a human chromosome 21-specific library will be described elsewhere. Briefly, human chromosome 21 was isolated from a male lymphoblast cell line having a normal karyotype. Crude chromosomes were extracted from M-phase cells and were stained with Hoechst 33258 and chromomycin A3 prior to separation of individual chromosomes by FISH analysis. Fluorescence in situ hybridization (FISH) analysis revealed a specific distribution of the repeat units among chromosomes. To the best of our knowledge, this 93 bp repetitive sequence is a new human STIR. In addition, the repeat was also found in the genomic DNA of other species including chimpanzee.

3. Results

3.1. Detection and mapping of the 93 bp repeat sequences

To study the structural characteristics of the region adjacent to the telomere of human chromosome 21, we isolated fosmid clone F50F5 and the P1 clone P1-11C1 from the chromosome 21-specific libraries using telomeric repeat and STS
markers. From the results of DNA sequencing, we identified 93 bp repeating sequences in which highly homologous but not identical unit sequences were tandemly repeated 10 times (Fig. 1a). The exact position of this 93 bp repeats is 7361–8103 nucleotides from the start point of the telomeric (TTAGGG)n repeats (Fig. 1b).

3.2. Characterization of the 93 bp repetitive sequences

The nucleotide sequences of the 93 bp repeating units are highly conserved (93–99%). No identical unit is present in our sequence, and the positions of nucleotide substitution seem to be quite limited within the repeating units (Fig. 1). To test if the 93 bp is unique to chromosome 21, we surveyed the currently available nucleic acid databases using the BLAST program at first; however, sequences showing clear similarity were detected in human chromosomes 10, 19, and 22. Interestingly, we also detected several sequences in the human EST database that have similarity to the 93 bp units. However, the 93 bp units do not have a repeating open reading frame. No oligopeptide having a corresponding amino acid sequence was found in the protein databases. Thus it is unlikely that the repeat structure either contains part of a gene cluster or encodes a repeating oligopeptide motif.

Since human genome sequencing is still in its early stages, the above observations led us to experimentally survey homologous sequences in the human genome by the FISH technique using the entire sequence or the repeat structure (Fig. 1b) as a probe. As shown in Fig. 2, we detected strong positive signals at the chromosomal loci 5pter, 7pter, 17qter, 19pter, 19qter, 20pter, 21qter, and 22qter. In addition, several weaker signals could be detected at the telomeric region of other chromosomes (Fig. 2). The variable signal strength might reflect different similarities or number of repeats between the repeat structures. The 93 bp repeat was detected in the genomic DNA of at least 10 different individuals (data not shown), indicating that the repeat is widely conserved in the human genome.

An interesting finding was that we detected strong hybridization signals at the position of 2q13. Because this position is known to be a remnant of an ancient chromosomal fusion event possibly coupled to the divergence of humans from other primates [10,11], we further tested the existence of the 93 bp repeat using genomic DNA of other species. As shown in Fig. 3, we detected the presence of similar 93 bp repeats in DNA obtained from two different chimpanzees, but no significant number of repeats was detected in mouse DNA using the same PCR primers as with humans (Figs. 1 and 3). Thus, we concluded that the 93 bp repeat is conserved not only in human populations but also among primates and in other species probably in a less prominent form. Since our present resources are limited, detailed analysis is left for future study.

4. Discussion

The results of FISH analysis showed a definite distribution of the 93 bp repeats in the telomeric region of several chromosomes, although their evolutionary relationship is unclear.

---

Fig. 1. Position and structure of the 93 bp repeat units. a: Position of the best known telomeric STS marker, HTRM1L1, which is located about 35 kb from the 21q telomere (H.-S. Park, unpublished observation). Fosmid clone F50F5 is the most telomeric clone isolated in this study. Positions of the PvuII site, and the start and end points of the 93 bp repeats are indicated. Nucleotide positions are numbered from the start position of the telomeric repeats. Dot matrix analysis demonstrated the presence of direct repeats between positions 7361 and 8103. b: The 93 bp repeat unit is repeated 10 times between nucleotide positions 7361 and 8103. Nucleotides identical to the first unit are replaced by dots. A single nucleotide gap is represented by an asterisk. Arrowheads represent two sets of primer oligonucleotides, Srep93 and Irep93.

Fig. 2. Distribution of the 93 bp repetitive sequences in the human chromosomes. a: Example of R-banded FISH analysis. Positive signals are shown as bright dots. One chromosome 2 is seen at the upper left corner. (b) Scores of positive signals are displayed to the right of the ideogram of each chromosome. Only signals detected at the same site on both chromatids were counted.
sequences in the subtelomeric region of the chromosomes found in the chimpanzee karyotype [10], which occurred a few times. The 93 bp repeat sequence (Figs. 1 and 2). This locus is located on chromosome band 2q13 was obtained using the probe specific for this region. The most telomeric clone, F50F5, may be a useful DNA substrate to study functions of telomeres. The entire human genome sequence and results of comparative genomic studies should shed further light on this issue in the near future.

References