Jpn J Human Genet 39, 421–425, 1994

Short Communication

DIRECT MAPPING OF THE HUMAN TATA BOX-BINDING PROTEIN (TBP) GENE TO 6q27 BY FLUORESCENCE *IN SITU* HYBRIDIZATION

Fumiko SAITO[1,*] Tohru YAMAMOTO,[2] Masami HORIKOSHI,[2] and Tatsuro IKEUCHI[1]

[1]*Department of Cytogenetics, Medical Research Institute, Tokyo Medical and Dental University, 1–5–45 Yushima, Bunkyo-ku, Tokyo 113, Japan*
[2]*Laboratory of Developmental Biology, Department of Cellular Biology, Institute of Molecular and Cellular Biosciences, The University of Tokyo, 1–1–1 Yayoi, Bunkyo-ku, Tokyo 113, Japan*

Summary TBP (TATA box-binding protein) participates in the expression of eukaryotic genes transcribed by RNA polymerases I, II, and III. Molecular cloning of human TBP revealed that the N-terminal region contains a polymorphic $(CAG)_n$ repeat. We report here the direct localization of human TBP gene to chromosome 6q2705→qter region by fluorescence *in situ* hybridization, using the cDNA clone with or without the $(CAG)_n$ repeat as a probe.

Key Words TBP gene, gene mapping, $(CAG)_n$ repeat, FISH

TBP (TATA box-binding protein), which was originally characterized as a TATA box-binding subunit of TFIID, plays an essential role in both basal and regulated transcription directed by RNA polymerases I, II, and III (Hernandez, 1993). The N-terminal region of human TBP contains a characteristic primary structure: a stretch of glutamine residue, that is encoded by trinucleotide $(CAG)_n$ repeat (Q-Run) (Hoffmann *et al.*, 1990). This trinucleotide $(CAG)_n$ repeat was found to be polymorphic in the normal population by the study using PCR (Polymeropoulos *et al.*, 1991). Recently, the polymorphic $(CAG)_n$ repeat and its unstable expansion have been found within the coding regions of the genes responsible for late onset neurodegenerative diseases, such as Huntington disease (HD), spinobulbar muscular atrophy (SBMA), spinocerebellar ataxia 1, 2 (SCA1, SCA2), and dentatorubral-pallidoluysian atrophy (DRPLA) (Sutherland and Richards, 1993;

Received August 29, 1994; Accepted September 16, 1994.
* To whom correspondence should be addressed.

Koide *et al.*, 1994; Nagafuchi *et al.*, 1994). The corresponding uninterrupted $(CAG)_n$ repeat (Gln stretch) in each gene is shown to contain from 11 to 36 Gln residues in the normal population (Sutherland and Richards, 1993). Since the N-terminal region of human TBP contains as many as 34 consecutive Gln residues, this region may also be subjected to unstable expansion, giving rise to a certain currently known disease, the locus of which has been mapped or not.

In the present study, we demonstrate the direct localization of the TBP gene to 6q2705→qter by fluorescence *in situ* hybridization (FISH). The chromosomal location and its comparison with known loci will be discussed.

MATERIALS AND METHODS

Prometaphase chromosome preparations for high-resolution R-band analysis were obtained from peripheral lymphocyte cultures of healthy donors by the method of methotrexate (0.1 μM) or thymidine (300 μg/ml) synchronization followed by 5-bromo-2′-deoxyuridine (30 μg/ml) incorporation (Takahashi *et al.*, 1991).

Two kinds of cDNA fragments were used as probes for FISH: one (1,258 bp) was prepared from a clone H10 containing open reading frame encoding the entire human TBP (335 a. a.) (Hoffmann *et al.*, 1990), and the other (718 bp) was prepared from a clone M14 that is deleted the region corresponding to the N-terminal amino acid sequences ranging from the 2nd to 155th. Both DNA fragments were labeled by nick translation with biotin-16-dUTP (Boehringer Mannheim). Hybridization conditions were the same as described previously (Takahashi *et al.*, 1991; Saito *et al.*, 1994). To detect efficiently the signals from the cDNA probes, the amplification procedure (Hori *et al.*, 1990) was applied. The hybridization signals were detected by the same method as described previously (Saito *et al.*, 1994).

RESULTS AND DISCUSSION

To determine directly the chromosomal location of the human TBP gene using fluorescence *in situ* hybridization (FISH), we initially hybridized the biotinylated cDNA fragment containing the entire coding region of human TBP derived from H10 clone to the R-banded prometaphase chromosomes of healthy donors. A total of 39 doublet signals were observed, and 15 of them (38.5%) were exclusively detected on the terminal region of chromosome 6q, 6q27, more precisely on the distal half of the band 6q27 (Fig. 1, a and b). The other signals were distributed randomly along the chromosomes, and there were no chromosome bands or subbands on which more than 2 fluorescent signals were localized. Although these findings suggest that the human TBP gene is located on chromosome 6q2705→qter region, the following problem remains to be solved.

Since we used a cDNA fragment containing the entire coding region of human

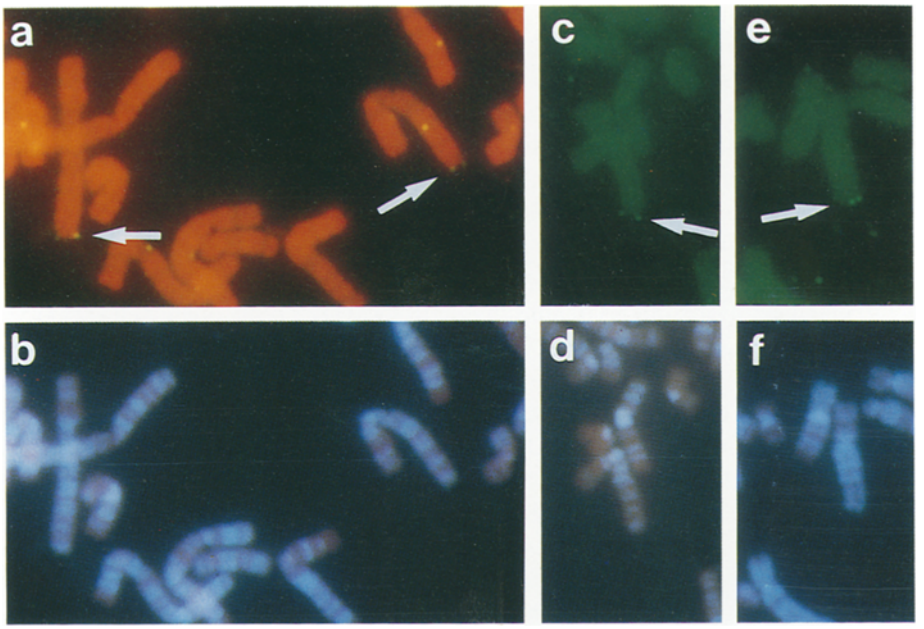

Fig. 1. FISH of biotinylated cDNA fragments derived from H10 (**a**) and M14 (**c** and **e**), on human chromosomes. **a**, **c**, and **e**: FITC signals (arrows) detected on the chromosomes using a Nikon B-2A (**a**) or B-2E (**c** and **e**) filter. **b**, **d**, and **f**: G-banding pattern on the same chromosomes as seen in **a**, **c**, and **e**, respectively, visualized using a Nikon UV-2A filter.

TBP as a probe, the possibility remains that the observed signals are due to hybridization of trinucleotide $(CAG)_n$ repeat in the TBP coding region to a similar repeat on unrelated chromosome DNA. Genes for some neurodegenerative diseases have been known to contain unstable CAG repeats and their chromosomal locations are: HD to 4p16.3, SCA1 to 6p21.3-p21.2, DRPLA to 12p, SCA2 to 12q23-q24, MJD to 14q24.3-q32.1, and SBMA to Xcen-q22 (O'Brien, 1993; Koide *et al.*, 1994). On these chromosomal regions, doublet signals were scarcely detected. However, still any region or gene containing the $(CAG)_n$ repeat structure may exist at the band 6q27.

To examine more concisely the above possibility, we prepared a cDNA fragment probe (M14) encoding only the C-terminal region of human TBP without any of the trinucleotide $(CAG)_n$ repeat. Although the length of this cDNA fragment was as short as 718 bp, we could still detect weak FITC signals through a B-2E filter on chromosome 6q27, with frequencies of 35% (15/43) for single and 16% (12/73) for doublet spots (Fig. 1, c–f). Since the cDNA fragment encoding human TBP lacking the $(CAG)_n$ repeat also gave signals on chromosome 6q27 as well as the cDNA fragment that contains the entire coding region did, we concluded that human TBP gene is unequivocally located on the chromosome region 6q2705→

qter. The results of our direct mapping using FISH were in agreement with that of indirect mapping by somatic hybrid cell lines (Polymeropoulos *et al.*, 1991), and linkage analyses (Imbert *et al.*, 1994), and clearly demonstrated the localization of this gene to the more specified region of chromosome 6.

It is difficult at present to speculate whether there might exist any candidate disease associated with the repeat expansion mutation of the TBP gene. So far, 4 hereditary disease loci have been mapped to the 6q27 region: cone dystrophy 3, retinal (COD3, 6q27), coronary atherosclerosis premature, with high Lpa (LPA, 6q26-q27), thrombosis, recurrent dysplasminogenemic type (PLG, 6q26-q27) and Best's macular dystrophy, vitelline, type 2 (VMD2, 6q25-qter) (HGM11). Among these genes, the COD3 and VMD2 genes have neither been cloned nor characterized. On the other hand, there are two kinds of t-complex genes (TCP1 and TCP10), murine T-locus homologs, among others mapped on the same chromosomal band region as TBP gene (O'Brien, 1993). Interestingly, the T gene in mice encodes a product with DNA binding activity which might control transcriptional regulation (Herrmann *et al.*, 1990; Kispert and Herrmann, 1993; Bollag *et al.*, 1994). This gene is likely to play a role in the development of all vertebrate organisms (Bollag *et al.*, 1994). Chromosome 6q27 might be one of the critical regions of human chromosomes, where the genes playing the essential role in proper embryonic development and cell differentiation are located.

Acknowledgments The authors thank Fumi Ishizaka for her kind technical help. This study was supported in part by the Special Coordination Funds of the Science and Technology Agency of the Japanese Government (T.I. & F.S.), Suzuken Memorial Foundation (M.H.), Shimadzu Science Foundation (M.H.) and Ministry of Education, Science and Culture of Japan (M.H. & T.Y.).

Note added in proof

After submitting this paper for publication, we have noticed a report (Purrello *et al.*, Genomics **22**: 94, 1994) with the same result as our mapping.

REFERENCES

Bollag RJ, Siegfried Z, Cebra-Thomas JA, Garvey N, Davison EM, Silver LM (1994): An ancient family of embryonically expressed mouse genes sharing a conserved protein motif with the T locus. Nature Genet **7**: 383–389

Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H (1990): Cloning of the T gene required in mesoderm formation in the mouse. Nature **343**: 617–622

Hernandez N (1993): TBP, a universal eukaryotic transcription factor. Genes Develop **4**: 1291–1308

HGM11 (1991): Eleventh International Workshop on Human Gene Mapping. Cytogenet Cell Genet **58**: 1–2200

Hoffmann A, Sinn E, Yamamoto T, Wang J, Roy A, Horikoshi M, Roeder RG (1990): Highly conserved core domain and unique N terminus with presumptive regulatory motifs in a human TATA factor (TFIID). Nature **346**: 387–390

Hori T, Takahashi E, Ayusawa D, Takeishi K, Kaneda S, Seno T (1990): Regional assignment

of the human thymidylate synthase (TS) gene to chromosome band 18q11.32 by nonisotopic *in situ* hybridization. Hum Genet **85**: 576–580

Imbert G, Trottier Y, Beckmann J, Mandel JL (1994): The gene for the TATA binding protein (TBP) that contains a highly polymorphic protein coding CAG repeat maps to 6q27. Genomics **21**: 667–668

Kispert A, Herrmann BG (1993): The *Brachyury* gene encodes a novel DNA binding protein. EMBO J **12**: 3211–3220

Koide R, Ikeuchi T, Onodera O, Tanaka H, Igarashi S, Endo K, Takahashi H, Kondo R, Ishikawa A, Hayashi T, Saito M, Tomoda A, Miike T, Naito H, Ikuta F, Tsuji S (1994): Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). Nature Genet **6**: 9–13

Nagafuchi S, Yanagisawa H, Sato K, Shirayama T, Ohsaki E, Bundo M, Takeda T, Tadokoro K, Kondo I, Murayama N, Tanaka Y, Kikushima H, Umino K, Kurosawa H, Furukawa T, Nihei K, Inoue T, Sano A, Komure O, Takahashi M, Yoshizawa T, Kanazawa I, Yamada M (1994): Dentatorubral and pallidoluysian atrophy expansion of an unstable CAG trinucleotide on chromosome 12p. Nature Genet **6**: 14–18

O'Brien SJ (eds) (1993): Human maps. In: Genetic maps, 6th edn., Cold Spring Harbor Press, Cold Spring Harbor, New York

Polymeropoulos MH, Rath DS, Xiao H, Merril CR (1991): Trinucleotide repeat polymorphism at the human transcription factor IID gene. Nucleic Acids Res **19**: 4307

Saito F, Sasaki S, Chepelinsky AB, Fushimi K, Marumo F, Ikeuchi T (1995): Human AQP2 and MIP genes, two members of the MIP family, map within chromosome band 12q13 on the basis of two-color FISH. Cytogenet Cell Genet **68**: 45–48

Sutherland GR, Richards RI (1993): Dynamic mutations on the move. J Med Genet **30**: 978–981

Takahashi E, Yamanouchi M, Tsuji H, Hitomi A, Meuth M, Hori T (1991): Chromosome mapping of the human cytidine-5′-triphosphate synthetase (CTPS) gene to band 1p34.1-p34.3 by fluorescence *in situ* hybridization. Hum Genet **88**: 119–121