Molecular Psychiatry (2000) 5, 32–38

www.nature.com/mp

IMMEDIATE COMMUNICATION

The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants

M Nakamura, S Ueno, A Sano and H Tanabe

The Department of Neuropsychiatry, Ehime University School of Medicine, Shitsukawa, Shigenobu, Onsen-gun, Ehime 791-0295, Japan

The serotonin transporter (5-HTT) gene is a promising candidate for introducing the heritability of interindividual variation in personality and the genetic susceptibility for various psychiatric diseases. Transcription of the gene is modulated by a common polymorphism in its upstream regulatory region (5-HTT gene-linked polymorphic region: 5-HTTLPR). The 5-HTTLPR consists of variation of the repetitive sequence containing GC-rich, 20-23-bp-long repeat elements. A deletion/insertion in the 5-HTTLPR was first reported to create a short (S) allele and a long (L) allele (14- and 16-repeats, respectively). Three other kinds of alleles (18-, 19- and 20-repeats) in addition to the S and L alleles in 5-HTTLPR have been reported. In the present study, we examined the 5-HTTLPR polymorphism in detail and identified ten novel sequence variants, concluding that the alleles reported as S and L are divided into four and six kinds of allelic variant, respectively. Subsequently, we developed a method for genotyping. The total number of alleles (14-A, 14-B, 14-C, 14-D, 15, 16-A, 16-B, 16-C, 16-D, 16-E, 16-F, 19, 20 and 22) in the 5-HTTLPR was 14 in our populations (Japanese: $n = 131$; Caucasian: $n = 74$) in the present study. In addition, a significant ethnic difference between Japanese and Caucasian populations was observed for distributions of alleles and genotypes ($P < 0.0001$ and $P < 0.0001$, respectively). Our results suggest that the analyses of the 5-HTTLPR should be revised by genotyping with a more complete subdivision of alleles. *Molecular Psychiatry* (2000) **5,** 32–38.

Keywords: serotonin transporter; functional polymorphism; novel alleles; ethnic difference

Introduction

Serotonin transporter (5-HTT) is involved in the presynaptic reuptake of serotonin to terminate and modulate serotonergic neurotransmission. The 5-HTT protein is the site of action of widely-used reuptake-inhibiting antidepressants such as selective serotonin-reuptake inhibitors and traditional tricyclic antidepressants. Therefore, a dysfunction of 5-HTT has been implicated in the etiology of psychiatric disorders such as mood, anxiety, obsessive-compulsive, and substance abuse disorders.

The human 5-HTT gene (SLC6A4) has been cloned and mapped on chromosome 17q11.1–q12.[1] Recently, a polymorphism has been identified in the region for transcriptional control of SLC6A4 (5-HTTLPR), which consists of different lengths of the repetitive sequence containing GC-rich, 20-23-bp-long repeat elements in the upstream regulatory region of the gene. A deletion/insertion in the 5-HTTLPR creates a short (S) allele and a long (L) allele (14- and 16-repeat alleles), which alters the promoter activity.[2–5] The S variant has been reported to be associated with lower basal and transcriptional efficiency of the SLC6A4, resulting in lower serotonin uptake activity when compared with the L variant.

The 5-HTTLPR polymorphism has been studied in various psychiatric disorders such as mood disorder,[4–6] schizophrenia,[7,8] panic disorder,[9] autistic disorder[10,11] and personality traits[3,12] resulting in positive or negative associations. Most of the studies have been performed using an easy PCR method with agarose gel electrophoresis to determine genotypes only for the S and L variants. Three other kinds of alleles (18-, 19- and 20-repeat[6,13,14]) as well as S and L alleles in 5-HTTLPR have also been studied previously.

In this study, we analyzed the 5-HTTLPR in detail and identified ten novel sequences, in addition to those previously reported, and studied their distribution in Japanese and Caucasian individuals.

Materials and methods

Subjects

Japanese subjects ($n = 131$) were selected from a population of unrelated Japanese residing in Ehime, Japan.

Correspondence: Dr A Sano, The Department of Neuropsychiatry, Ehime University School of Medicine, Shitsukawa, Shigenobu, Onsen-gun, Ehime 791-0295, Japan. E-mail: sano@m.ehime-u.ac.jp
Received 30 August 1999; revised and accepted 6 October 1999

All subjects gave written informed consent to participate in the study. Genomic DNA was prepared from peripheral blood leukocytes according to the standard method. Unrelated Caucasian samples ($n = 76$) were chosen from the genomic DNAs of CEPH families.

Genotyping to generate the S- and L- fragments
Polymerase chain reaction (PCR)-amplification was carried out with the primers, 5′-GGCGTTGCCGC TCTGAATGC-3′ (stpr5-2) and 5′-GAGGGACT GAGCTGGACAACCCAC-3′ (stpr3-2), in a total volume of 12.5 μl solution containing 100 ng genomic DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1 mM $MgCl_2$, 200 μM dNTPs (dGTP/7-deaza-2′-dGTP = 1/1), 5% dimethyl sulfoxide, 0.2 μM of each primer, and 0.5 U of Taq polymerase (Takara, Tokyo, Japan) for 35 cycles (94°C for 30 s, 61°C for 30 s, 72°C for 1 min). The PCR products were then analyzed in 1.5–2% agarose gel stained with ethidium bromide.

SSCP and heteroduplex analysis and molecular weight determination by denatured PAGE
Single-stranded conformational polymorphism (SSCP) analysis was performed using the DNA fragments amplified from genomic DNA of unrelated individuals. The PCR-amplification for 5-HTTLPR was performed with the primers, 5′-GCAACCTCCCAGCAACTCC-3′ (sert71F) and 5′-GGGATGCGGGGGAATACTG-3′ (sert384R), in a total volume of 12.5 μl containing 30–100 ng genomic DNA, 1.25 μl of 10 × Expand HF Buffer (Boehringer Mannheim, Mannheim, Germany), 1 mM $MgCl_2$, 200 μM dNTPs, 1 μCi [α-^{32}P]dCTP, 5% dimethyl sulfoxide, 0.2 μM of each primer, and 0.7 U of Expand™ High Fidelity PCR system enzyme mix (Boehringer Mannheim) for 35 cycles (94°C for 15 s, 58°C for 30 s and 68°C for 30 s). The PCR products were denatured and run onto a 6% non-denaturing polyacrylamide gel at 25°C/40 W either with or without 5% glycerol in the gel (for SSCP), or onto a 6% denaturing polyacrylamide gel including 4 M urea at 50°C/45 W (for molecular weight determination). Gels were dried and exposed to imaging plates (Fuji Xerox, Tokyo, Japan); the photostimulated luminescence of the bands was analyzed by BAS1000 system (Fuji Xerox).

Sequence analysis
Sequencing the polymorphic variants detected by the SSCP analysis was performed as follows: The fragments were PCR-amplified in a total volume of 50 μl solution containing 100 ng genomic DNA, 5 μl of 10 × Expand HF Buffer (Boehringer Mannheim), 1 mM $MgCl_2$, 200 μM dNTPs, 0.2 μM of each primer (stpr5-2 and stpr3-2), and 1.4 U of Expand™ High Fidelity PCR system enzyme mix (Boehringer Mannheim) for 30 or 35 cycles (94°C for 15 s, 61°C for 30 s and 68°C for 45 s). The PCR products were electrophoresed on 2% agarose gels in 0.5 × TBE, stained with ethidium bromide, and appropriately sized bands were excised. DNA was extracted from the gel slices using Gene Clean II kit (Bio101, Vista, CA, USA). The genomic DNA samples for amplification were selected to be heterozygous for different repeat alleles, which enabled the amplified products to be separated by agarose gel electrophoresis and by extraction. The extracted products were directly sequenced. The PCR product amplified from genomic DNA of heterozygous 14-A/14-C was cloned into vectors (pT7Blue T-vector Kit, Novagen, Madison, WI, USA), only in the case of 14-C. Direct sequencing was performed using BigDye Terminator Cycle Sequencing FS Ready Reaction Kit (PE Applied Biosystems, Foster City, CA, USA). Sequencing fragments were analyzed by ABI PRISM 310 Genetic analyzer (PE Applied Biosystems). Nucleotide sequences of the cloned inserts were determined using Bca BEST™ Dideoxy Sequencing Kit (Takara) followed by electrophoresis on 6% denaturing polyacrylamide gels.

PCR-restriction fragment length polymorphism (RFLP) analysis
PCR-amplification was performed for 5-HTTLPR with the primers, stpr5-2 and stpr3-2, in a total volume of 12.5 μl solution containing 100 ng genomic DNA, 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1 mM $MgCl_2$, 200 μM dNTPs (dGTP/7-deaza-2′-dGTP = 1/1), 5% dimethyl sulfoxide, 0.2 μM of each primer and 0.5 U of Taq polymerase (Takara) for 35 cycles (94°C for 30 s, 61°C for 30 s, 72°C for 1 min). Each of the resultant PCR products was digested overnight by each 10 unit of *Msp*I or *Apa*LI (New England Biolabs, MA, USA), with additional amounts of dithiothreitol and $MgCl_2$ added to the PCR mixture to make final concentrations of 1 and 10 mM, respectively. The digested PCR products were electrophoresed on a 3% NuSeive® 3:1 agarose gel (FMC, Rockland, ME, USA) stained with ethidium bromide.

PCR-RFLP analysis using mutagenic primers
For discrimination of the same length of 16-repeat variants (16-A, 16-E and 16-F), PCR-RFLP was performed using mutagenesis primers to introduce the polymorphic restriction site. PCR-amplifications were performed using the primers to determine 16-E, 5′-CCCCCCAGCATCTCCCCTGGA-3′ (5htt-185F) and 5′-GCTTGCCCGTGGTTGTCCAGCTCAG-3′ (5htt-499R); and 16-F, 5′-CTCGCGGCATCCCCCCTGGA-3′ (5htt-143F) and 5htt-499R. The forward primers, 5htt-185F and 5htt-143F, introduce base alterations (C → G, the underlined of each primer) into the PCR products, which creates an artificial *Ava*II recognition site (GGACC). The final concentrations for PCR reactions were 100 ng genomic DNA; 10 mM Tris-HCl (pH 8.3), 50 mM KCl; 1 mM $MgCl_2$; 200 μM dNTPs (dGTP/7-deaza-2′-dGTP = 1/1); 5% dimethyl sulfoxide; 0.2 μM of each primer; and 0.5 U of Taq polymerase (Takara) for 35 cycles (94°C for 30 s, 61°C for 30 s, 72°C for 1 min). Subsequently, PCR products were digested with 10 units of *Ava*II and electrophoresed on a 3% NuSeive® 3:1 agarose gel (FMC) stained with ethidium bromide.

Statistical analysis

For statistical evaluation of Hardy–Weinberg equilibrium and the difference between groups of Caucasians and Japanese with respect to the allelic and genotypic frequencies, an exact test was performed using the program Arlequin ver 1.1.[15] The exact test is analogous to Fisher's exact test on a 2×2 contingency table, but extends to a contingency table of arbitrary size.[16,17]

Nucleotide sequences

The nucleotide sequence data reported in this paper will appear in the DDBJ/EMBL/GenBank nucleotide sequence database with accession numbers AB031247–AB031259.

Results

Identification of 15- and 22-repeat alleles

We observed six kinds of alleles in agarose gel separation when we performed genotyping to generate S- and L-fragments (Figure 1a). Subsequent sequencing analyses showed that the alleles consisted of 14-, 15-, 16-, 19-, 20-, and 22-repetitive elements (Figure 2). The 15- and 22- repeat alleles have not been described up to the present, while the 19- and 20-repeat variants had the identical sequences as those reported by Kunugi *et al*.[6] The most frequent 14-repeat allele as mentioned later was named '14-A' in this report. All repeat elements of 14-A showed sequences different from one

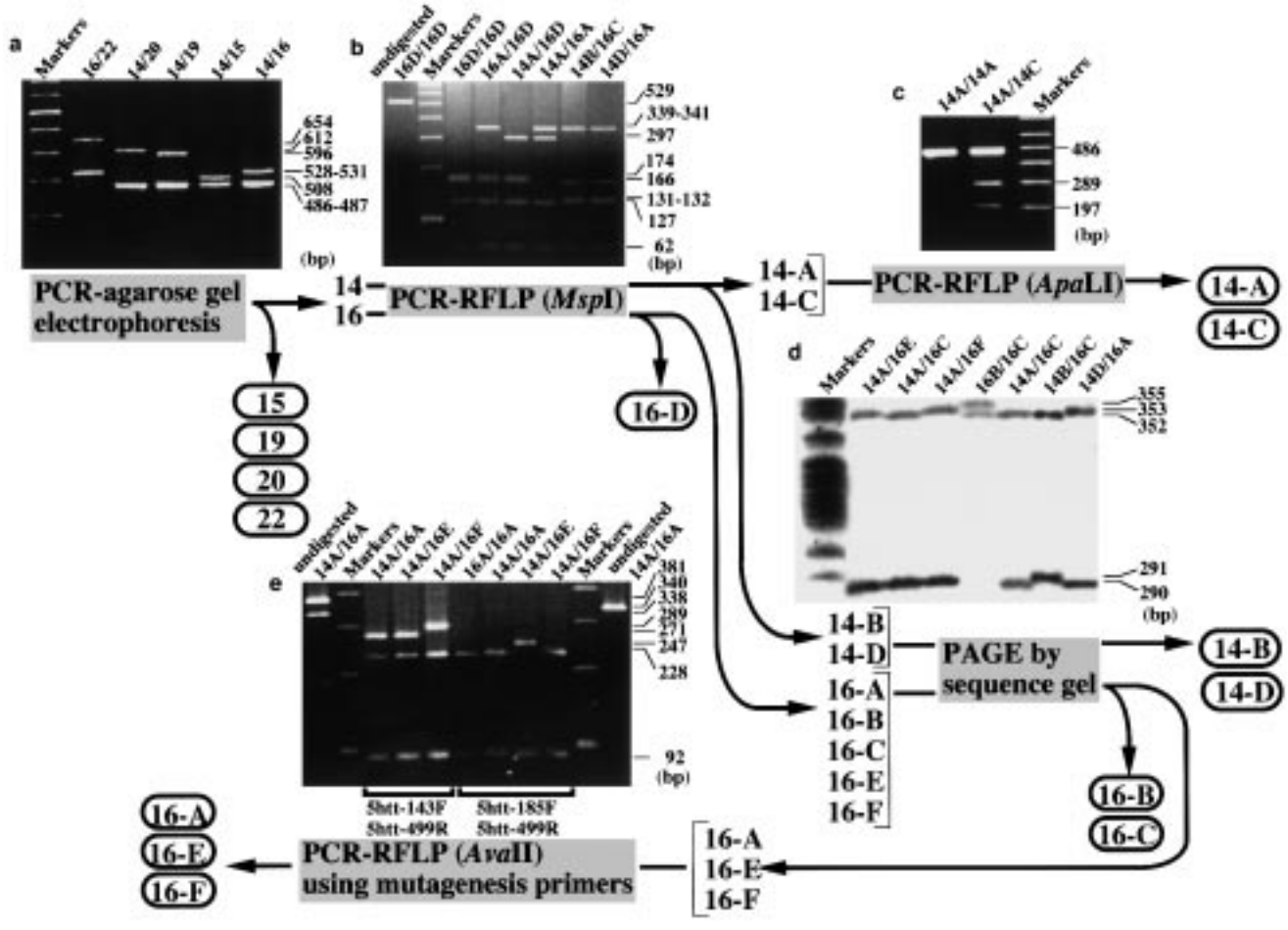

Figure 1 Flow chart for genotyping. Names of the alleles determined after analyses are shown in circles. (a) Agarose gel electrophoresis after the simple non-labeled PCR. 15-, 19-, 20- and 22-repeat alleles were identified in this step. (b) Restriction fragment length polymorphism (RFLP) analysis, with the enzyme *Msp*I, of PCR products carrying 14-repeat or 16-repeat alleles. The 16-D allele was identified in this step (174- and 166-bp fragments). The 14-repeat alleles are subdivided into two groups. One includes 14-A and 14-C, the other consists of 14-B and 14-D (166-bp fragment). (c) Restriction analysis (*Apa*LI) of PCR products carrying the 14-A or 14-C allele. PCR products carrying the 14-C allele were digested into 289-bp and 197-bp fragments. (d) Molecular weight determination by denatured PAGE. The radiolabeled PCR amplification for samples unidentified in (a)–(d) steps were performed with the primers sert71F and sert384R. The PCR products were denatured and run onto 6% denaturing polyacrylamide gel. The 16-B, 16-C, 14-B and 14-D alleles were identified (355, 352, 291, 290 bp, respectively). (e) PCR-RFLPs analysis using mutagenic primers. PCR amplifications for samples carrying 16-A, 16-E and 16-F alleles were performed with the primers, 5htt-143F and 5htt-499R or 5htt-185F and 5htt-499R. The forward primers, 5htt-185F and 5htt-143F, introduce base alterations into the PCR products, which creates an artificial *Ava*II recognition site. The 16-F allele was identified using the primers, 5htt-143F and 5htt-499R (289 bp), and the 16-E allele was identified using the primers, 5htt-185F and 5htt-499R (247 bp). The remains were 16-A alleles.

another, and they were named in order from the 5′-end as α, β, ..., and ξ as in Figure 2. The 15-repeat allele had an extra repeat element of 23 bp (ρ) which is similar to the repeat element ε. The 22-repeat allele had a tandem repeat containing pairs of ζ and η repeat elements.

SSCP-heteroduplex analyses

SSCP analyses revealed that the 14- and 16-repeat alleles in our populations were subdivided into four and six variants, respectively (Figure 3a). Since the differences in band shifts of the heteroduplexes facilitated the distinction between the variants, we added the 14-repeat allele to make artificial heterozygous 14-repeat/16-repeat when we analyzed the homozygous 16-repeat/16-repeat (Figure 3a).

Variants of 16-repeat alleles

Denatured PAGE analysis demonstrated that the 16-repeat alleles consisted of three kinds (352 bp, 353 bp, and 355 bp) (Figure 1d). In the SSCP analysis of Japanese samples, we observed four different kinds of band shifts for the 16-repeat alleles (Figure 1a). From the results of the sequence analyses, these alleles possessed similar but distinct arrays of repeat elements at the position of ζ element of the 14-A allele (16-A, 16-B, 16-C, and 16-D, Figure 2). In the SSCP analysis of the Caucasian samples, we found two additional kinds of band shifts (Figure 3b). The sequences of these alleles had also different arrays of repeat elements at the position of ζ element of the 14-A allele (16-E and 16-F, Figure 2). The sequence of the repeat elements in 16-A was exactly identical to the sequence reported by Kunugi *et al*[6], and it was almost identical to the L variant reported by Heils *et al*[2] (EMBL/GenBank accession No. X76753) except for three minor changes (Figure 2).

Variants of 14-repeat alleles

In the SSCP analysis of Japanese samples, we observed three kinds of band shifts for the 14-repeat alleles (14-

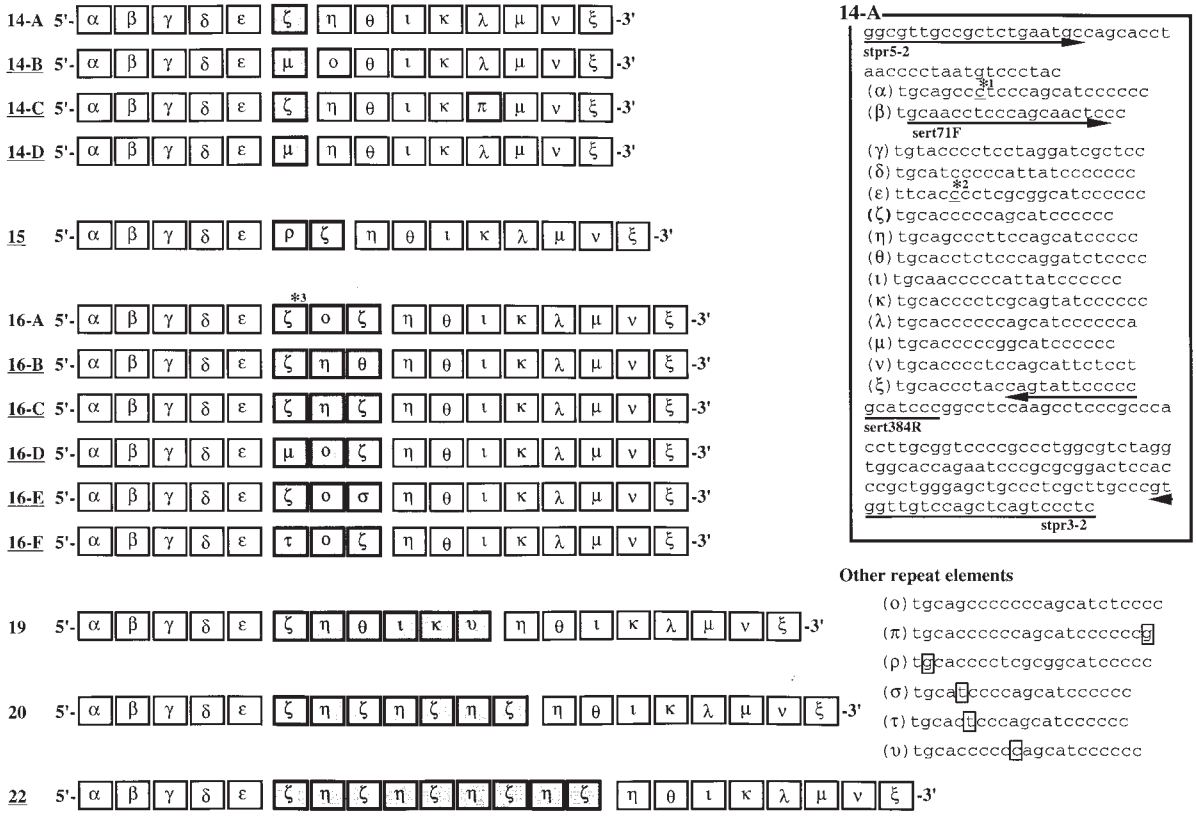

Figure 2 Nucleotide sequences and repeat architectures of 5-HTTLPR alleles. Names of the alleles newly identified in the present study are underlined. The sequence of the 14-A allele is shown in the upper right corner in the inset. The locations and orientations of the oligonucleotides used as the primers for PCR are shown by arrows. The repeat elements of 14-A were slightly different from each other, and they were named in order from the 5′-end as α, β, ..., and ξ. The elements shown in dark boxes are inserted in the place of ζ element of the 14-A allele. Comparing our sequences to that reported by Heils *et al*,[2] three minor changes shown by asterisks 1–3 were found. In the sequence by Heils *et al*, the nucleotide 'c' shown by * 1 was not present, the nucleotide 'c' shown by * 2 was 't', and the ζ element in the 16-A allele shown by * 3 was υ. In the lower left corner, the repetitive elements other than those in the 14-A are shown. The nucleotides in boxes are different from the nucleotides in similar repeat elements. The element π is different from λ by substitution of 'a' to 'g'. The element ρ is different from ε by substitution of 't' to 'g'. The element σ is different from ζ by substitution of 'c' to 't'. The element τ is different from ζ by substitution of 'c' to 't'. The element υ is different from ζ by insertion of 'c'.

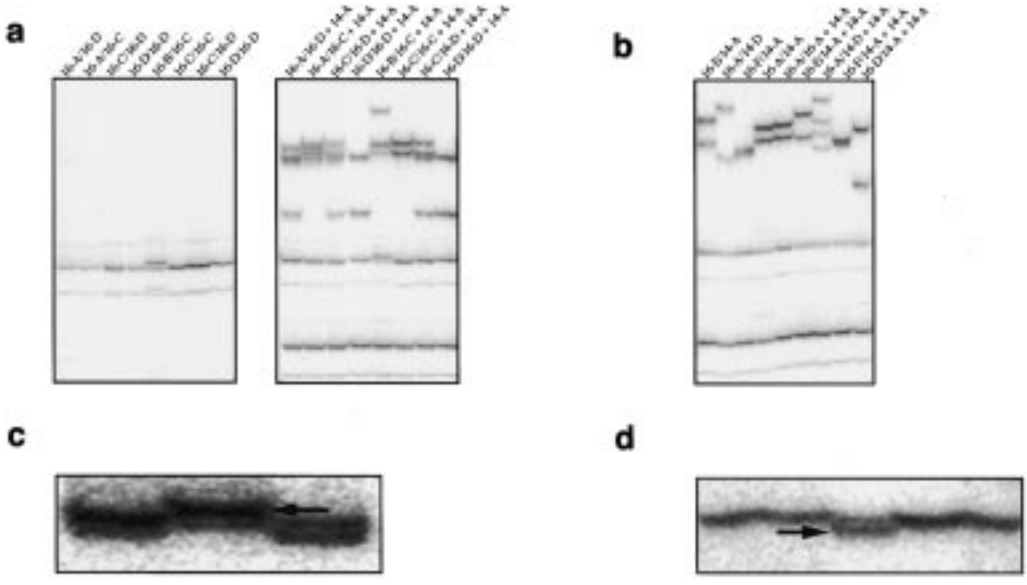

Figure 3 SSCP and heteroduplex analyses of 5-HTTLPR. (a) The SSCP analysis of 16-repeat/16-repeat alleles is shown in the left panel. The right panel shows the electrophoretic patterns of the artificially produced heteroduplexes of the samples of 16-repeat/16-repeat mixed with 14-A. The electrophoresis was performed in 6% non-denaturing polyacrylamide gels at 25°C/40 W without glycerol. (b) The heteroduplex bands are quite useful to discriminate alleles in the SSCP analysis. They were reproducible and artificially produced with the addition of certain alleles. (c) The arrow shows the band shift of the 14-B allele. (d) The arrow shows the band shift of the 14-C allele.

A, 14-B, 14-C; Figure 3c, 3d). In Caucasians, an additional 14-repeat allele (14-D) was found (Figure 3b). In total, four kinds of 14-repeat alleles were identified by sequence determination (Figure 2). The sequence of the repeat elements of the 14-A was exactly identical to the sequence reported by Kunugi *et al*[6], and it was almost the same as the S variant reported by Heils *et al*[2] (EMBL/GenBank accession No. X76753) except for two minor changes (Figure 2).

Developing a method for genotyping

Figure 1 shows a flow chart for genotyping analysis. First we performed PCR-amplification with the primers stpr5-2 and stpr3-2 without the use of radioisotopes, then the amplified products were separated in 1.5–2% agarose gel in order to determine 14-, 15-, 16-, 19-, 20- and 22-repeat alleles (486–487 bp, 508 bp, 528–531 bp, 596 bp, 612 bp and 654 bp, respectively, Figure 1a). To identify 16-D from the 16-repeat alleles, we performed PCR-RFLP with the restriction enzyme, *Msp*I; the 16-D allele produces 166- and 174-bp fragments whereas other 16-repeat alleles have 339–341 bp fragments (Figure 1b). To differentiate 14-B and 14-D from other 14-repeat alleles, we performed PCR-RFLP with the restriction enzyme, *Msp*I. The 14-B and 14-D alleles produce 166-bp and 131–132 bp fragments (14-B, 132 bp; 14-D, 131 bp) whereas other 14-repeat alleles produce 297-bp fragments (Figure 1b). To subdivide the remaining 14-repeat alleles into 14-A and 14-C, we performed PCR-RFLP with the restriction enzyme, *Apa*LI. The 14-C allele produces 289-bp and 197-bp fragments whereas the 14-A allele produces 486 bp fragments (Figure 1c). To differentiate 16-B and 16-C from other 16-repeat alleles and to differentiate 14-B from other 14-D, we performed radiolabeled PCR with the primers sert71F and sert384R followed by fragment lengths determination by denatured PAGE. The groups of 16-A and 16-E and 16-F, and 16-B, and 16-C alleles were 353 bp, 355 bp, and 352 bp, respectively. The 14-B and 14-D were 291 bp and 290 bp, respectively (Figure 1d). To differentiate between 16-A, 16-E and 16-F, PCR-RFLPs were performed using mutagenesis primers to introduce the *Ava*II recognition site (Figure 1e). To detect the 16-E allele, PCR-amplification with primer, 5htt-185F and 5htt-499R, was performed and the resultant PCR products were digested by *Ava*II. 16-E produces 247-bp fragments whereas both 16-A and 16-F produce 228-bp and 19-bp fragments. To detect the 16-F allele, PCR-amplification with primers 5htt-143F and 5htt-499R was performed and the PCR products were digested by *Ava*II. 16-F produces 289-bp fragments whereas 16-A and 16-E produce 271-bp and 18-bp fragments.

Distribution of alleles and genotypes in Japanese and Caucasians

Distributions of genotypes and alleles are shown in Table 1. The genotype frequencies were in Hardy–Weinberg equilibrium (Japanese: $P = 0.812$; Caucasian: $P = 0.861$). There were a total of fourteen different alleles (14-A, 14-B, 14-C, 14-D, 15, 16-A, 16-B, 16-C, 16-D, 16-E, 16-F, 19, 20 and 22) in the 5-HTTLPR in our samples (Japanese: $n = 131$; Caucasian: $n = 74$). In Japanese samples, a total number of eleven types of alleles were found (14-A, 14-B, 14-C, 15, 16-A, 16-B, 16-C, 16-D, 19, 20 and 22) (Table 1). In Caucasian samples, six types of alleles were found (14-A, 14-D, 16-A, 16-D, 16-E and 16-F). There were significant dif-

Table 1 The allelic and genotypic frequencies of 5-HTTLPR in Caucasian and Japanese populations

Allelic frequency

Allele	*Caucasian (n = 148)*	*Japanese (n = 262)*
14-A	66 (44.6)	207 (79.0)
14-B	0 (0.0)	1 (0.4)
14-C	0 (0.0)	1 (0.4)
14-D	1 (0.7)	0 (0.0)
16-A	72 (48.6)	18 (6.9)
16-B	0 (0.0)	3 (1.1)
16-C	0 (0.0)	13 (5.0)
16-D	7 (4.7)	11 (4.2)
16-E	1 (0.7)	0 (0.0)
16-F	1 (0.7)	0 (0.0)
15	0 (0.0)	2 (0.8)
19	0 (0.0)	1 (0.4)
20	0 (0.0)	4 (1.5)
22	0 (0.0)	1 (0.4)

The percentage of each allele for the samples is listed in parentheses.

Genotypic frequency

Genotype	*Caucasian (n = 74)*	*Japanese (n = 131)*
14-A/14-A	15 (20.3)	80 (61.1)
14-A/14-C	0 (0.0)	1 (0.8)
14-A/16-A	30 (40.5)	15 (11.5)
14-A/16-B	0 (0.0)	3 (2.3)
14-A/16-C	0 (0.0)	11 (8.4)
14-A/16-D	3 (4.1)	11 (8.4)
14-A/16-E	1 (1.4)	0 (0.0)
14-A/16-F	1 (1.4)	0 (0.0)
14-B/16-C	0 (0.0)	1 (0.8)
14-D/16-A	1 (1.4)	0 (0.0)
16-A/16-A	19 (25.7)	0 (0.0)
16-A/16-C	0 (0.0)	1 (0.8)
16-A/16-D	4 (5.4)	0 (0.0)
14-A/15	0 (0.0)	2 (1.5)
14-A/19	0 (0.0)	1 (0.8)
14-A/20	0 (0.0)	3 (2.3)
16-A/20	0 (0.0)	1 (0.8)
16-A/22	0 (0.0)	1 (0.8)

The percentage of each genotype for the samples is listed in parentheses.

ferences in both allelic and genotypic frequencies between Japanese and Caucasian populations (allelic frequency: $P < 0.0001$; genotypic frequency: $P < 0.0001$).

Discussion

We identified ten novel variants in 5-HTTLPR and found that the 5-HTTLPR consists of a total of 14 kinds of alleles in Japanese and Caucasian populations. Most of the differences in the 5-HTTLPR were explained by the differences in the arrays of inserted repeat elements in the place of the ζ element of the 14-A allele (Figure 2). This suggests that the sequence around the ζ element may represent a recombination hotspot. Heils *et al* pointed out the presence of a sequence of 'hot spot' for deletion mutagenesis (TGCAGCC) in the η repeat element.[2,18,19] This TGCAGCC-sequence appears repeatedly at the inserted sequences of the various alleles identified in this report, and only the ε element begins with sequence TTCA, instead of the TGCA sequence that begins all other elements.

Similar extreme variation is present in the 5′-flanking region of the human insulin gene. This gene contains a section of extensive length and sequence heterogeneity generated by variation in the number and arrangement of elements of a family of oligonucleotides whose consensus sequence is ACAGGGGTGTGGG.[20] Variation in this GC-rich variable number of tandem repeat (VNTR) is directly associated with the expression of insulin and insulin-like growth factor genes. The presence of hotspot sequences and GC-richness in the VNTR might explain the extreme variation in the 5-HTTLPR. Another possible reason for the alternating nature of the longer repeats is some form of slippage during DNA replication.

The 20-repeat allele in the present study was identical to the one reported by Kunugi *et al*.[6] Delbrück *et al* reported another 20-repeat variant which differs in one repetitive element from our 20-repeat variant at the inserted sequence in the place of the ζ element of the 14-A allele.[21] They found their 20-repeat allele only in the African population and not in Caucasian and Asian populations.

Ethnic differences in allelic frequencies of the 5-HTTLPR have been described elsewhere, mainly by division of the S and L alleles.[6,9,14,21] In the present study, significant ethnic differences between Japanese and Caucasian were observed for the distributions of allele and genotype subdivided into a total of 14 allelic variants. Concerning the 14-repeat alleles, the frequencies of 14-B, 14-C and 14-D were quite rare (14-B and 14-C were not found in Caucasians, 14-D was not found in Japanese) and the 14-A was by far the most common among 14-repeat alleles in both populations. Concerning the 16-repeat alleles, Japanese had relatively frequent four variants (16A, 16-B, 16-C and 16-D; 16-E and 16-F were not found), while Caucasians had mainly two variants (16-A and 16-D; 16-B and 16-C were not found) and also rare 16-E and 16-F alleles. The 15-, 19-, 20- and 22-repeat alleles were rare and found only in the Japanese population. Significant ethnic difference in the 5-HTTLPR polymorphism may lead to specific psychiatric disorders or personality. This has an important bearing on the potential for the large SNP projects to reveal variation that may be globally useful.

Although we developed a complete method for genotyping the 5-HTTLPR polymorphism into 14 variants in the present study (Figure 1), further studies should be performed. The 5-HTTLPR is a highly polymorphic region, and there is a possibility of the presence of different variants in other ethnically dissimilar popu-

lations. The SSCP analysis described here is highly reproducible and very useful, especially when band shifts of the heteroduplexes are carefully observed.

The 5-HTTLPR polymorphism has been studied in several psychiatric disorders and personality conditions. Many positive associations have been reported.[3,4,6,9,10] Other studies report non-replication.[5,7,8,11] Most of the studies have been performed using a simple PCR method with agarose gel electrophoresis to determine genotypes for the S and L variants. Ambiguous associations between diseases and 5-HTTLPR may be due to analyses discriminating only the S and L types.

The 5-HTTLPR region serves as a promoter for the 5-HTT gene and it was shown that the long variant possesses higher transcriptional activity than the short variant.[2–5] We are aware of the importance of the functional characterization of the alleles shown in the present study and plan functional analyses.

To conclude, our results suggest that analysis of the 5-HTTLPR should be revised by genotyping with a more expansive subdivision of alleles.

Acknowledgements

We would like to express our thanks to the subjects who participated in this study, to Miss Y Maruki for her technical assistance, and to Dr Y Nakamura of the University of Tokyo for providing genomic DNAs of the CEPH family members. This work was supported in part by a Grant-in-Aid for Co-operative Research from the Ministry of Education, Science, Sports and Culture of Japan.

References

1 Ramamoorthy S, Bauman AL, Moore KR, Han H, Yang-Feng T, Cheng A *et al*. Antidepressant- and cocaine-sensitive human serotonin transporter: molecular cloning, expression, and chromosomal localization. *Proc Natl Acad Sci USA* 1993; **90**: 2542–2546.
2 Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D *et al*. Allelic variation of human serotonin transporter gene expression. *J Neurochem* 1996; **66**: 1–4.
3 Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S *et al*. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science* 1996; **274**: 1527–1531.
4 Collier DA, Stöber G, Li T, Heils A, Catalan M, Di Bella D *et al*. A promoter of the serotonin gene: possible role in susceptibility to affective disorder. *Mol Psychiatry* 1996; **1**: 453–460.
5 Mortensen OV, Thomassen M, Larsen MB, Whittemore SR, Wiborg O. Functional analysis of a novel human serotonin transporter gene promoter in immortalized raphe cells. *Brain Res Mol Brain Res* 1999; **68**: 141–148.
6 Kunugi H, Hattori M, Kato T, Tatsumi M, Saki T, Sasaki T *et al*. Serotonin transporter gene polymorphisms: ethnic difference and possible association with bipolar affective disorder. *Mol Psychiatry* 1997; **2**: 457–462.
7 Malhotra AK, Goldman D, Mazzanti C, Clifton A, Breier A, Pickar D. A functional serotonin transporter (5-HTT) polymorphism is associated with psychosis in neuroleptic-free schizophrenics. *Mol Psychiatry* 1998; **3**: 328–332.
8 Mendes de Oliveira JR, Otto PA, Vallada H, Lauriano V, Elkis H, Lafer B *et al*. Analysis of a novel functional polymorphism within the promoter region of the serotonin transporter gene (5-HTT) in Brazilian patients affected by bipolar disorder and schizophrenia. *Am J Med Genet* 1998; **81**: 225–227.
9 Matsushita S, Muramatsu T, Kimura M, Shirakawa O, Mita T, Nakai T *et al*. Serotonin transporter gene regulatory region polymorphism and panic disorder. *Mol Psychiatry* 1997; **2**: 390–392.
10 Cook EH, Courchesne R, Lord C, Cox NJ, Yan S, Lincoln A *et al*. Evidence of linkage between the serotonin transporter and autistic disorder. *Mol Psychiatry* 1997; **2**: 247–250.
11 Klauck SM, Poustka F, Benner A, Lesch KP, Poustka A. Serotonin transporter (5-HTT) gene variants associated with autism? *Hum Mol Genet* 1997; **6**: 2233–2238.
12 Flory JD, Manuck SB, Ferrell RE, Dent KM, Peters DG, Muldoon MF. Neuroticism is not associated with the serotonin transporter (5-HTTLPR) polymorphism. *Mol Psychiatry* 1999; **4**: 93–96.
13 Michaelovsky E, Frisch A, Rockah R, Peleg L, Magal N, Shohat M *et al*. A novel allele in the promoter region of the human serotonin transporter gene. *Mol Psychiatry* 1999; **4**: 97–99.
14 Delbrück SJ, Wendel B, Grunewald I, Sander T, Morris-Rosendahl D, Crocq MA *et al*. A novel allelic variant of the human serotonin transporter gene regulatory polymorphism. *Cytogenet Cell Genet* 1997; **79**: 214–220.
15 Stefan S, Jean-Marc K, David R, Laurent E. *Arlequin: a Software for Population Genetic Data Analysis*. Genetics and Biometry Laboratory, University of Geneva: Switzerland, 1997.
16 Guo SW, Thompson EA. Performing the exact test of Hardy–Weinberg proportion for multiple alleles. *Biometrics* 1992; **48**: 361–372.
17 Raymond M, Rousset F. An exact test for population differentiation. *Evolution* 1995; **49**: 1280–1283.
18 Cooper DN, Krawczak M, Antonorakis SE. *The Metabolic and Molecular Bases of Inherited Disease*, 7th edn. McGraw-Hill: New York, 1995, pp 259–291.
19 Cooper DN, Krawczak M. *Human Gene Mutation*. Bios Scientific: Oxford, 1993.
20 Bell GI, Selby MJ, Rutter WJ. The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 1982; **295**: 286–292.
21 Gelernter J, Cubells JF, Kidd JR, Pakstis AJ, Kidd KK. Population studies of polymorphisms of the serotonin transporter protein gene. *Am J Med Genet* 1999; **88**: 61–66.