J Hum Genet (2002) 47:213–215

SHORT COMMUNICATION

Shoko Higuchi · Yusuke Nakamura · Susumu Saito

Characterization of a VNTR polymorphism in the coding region of the *CEL* gene

Received: December 26, 2001 / Accepted: January 30, 2002

Abstract Human carboxyl ester lipase (CEL) secreted by the pancreas into the duodenum is a glycoprotein playing an essential role in the intestinal processing of cholesterol and lipid-soluble vitamins. The gene encoding CEL was known to contain a tandemly repeated sequence of the 11-amino-acid motif in the C-terminal region. We characterized its polymorphic features and found that there are five different alleles in Japanese populations and six in Caucasians. The allele containing 16 repeats is the most common in both populations. Although the distribution of the alleles seemed to be different in the two populations, the difference was not statistically significant. This polymorphism may influence the function of this enzyme and be a useful genetic marker to study diseases associated with cholesterol absorption.

Key words Variable number of tandem repeats (VNTR) · Carboxyl-ester lipase (CEL) · Bile-salt-dependent lipase (BSDL) · PEST sequence · Chromosome 9

Introduction

Carboxyl ester lipase (CEL, EC3.1.1), also called bile-salt-dependent lipase, is a major component of pancreatic juice (Lombardo et al. 1978). It plays an important role in hydrolysis and absorption of cholesterol and lipid-soluble vitamin esters (Lombardo et al. 1980). The *CEL* gene contains a GC-rich region in the 3' portion, which encodes 16 repeats of an 11-amino-acid unit (Nilsson et al. 1990). Southern analysis indicated the CEL locus to be highly polymorphic and likely to contain a hypervariable region (Taylor et al. 1991).

S. Higuchi · Y. Nakamura · S. Saito (✉)
Laboratory for Genotyping, SNP Research Center, Institute of Physical and Chemical Research, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Tel. +81-3-5449-5716; Fax +81-3-5449-5433
e-mail: s-saito@ims.u-tokyo.ac.jp

The consensus unit of this 11-amino-acid tandem-repeated sequence contains three proline (P) residues, one glutamic acid (E) residue, one serine (S) residue, and one threonine (T) residue, and is referred to as a PEST sequence, which acts as a signal for rapid degradation of proteins (Rogers et al. 1986). PEST regions are considered to include a potential site for *O*-linked glycosylation (Mas et al. 1993). *O*-glycosylation of these sequences was suggested to mask the PEST domain and to be required for CEL secretion (Bruneau et al. 1997).

We here describe the variable number of tandem repeats (VNTR) feature of this tandemly repeated region and the characterization of allele frequencies in the Japanese and Caucasian populations.

Subjects and methods

DNA sources. Genomic DNAs were isolated from peripheral leukocytes of 46 unrelated Japanese individuals by the standard phenol/chloroform extraction method. A written informed consent for DNA analysis was obtained from all individuals. In addition, DNA samples from 31 unrelated Caucasian individuals used in this study were obtained from Centre d'Etude du Polymorphisme Humain.

PCR. On the basis of sequence information from GenBank (accession number AF072711.1), we designed polymerase chain reaction (PCR) primers to specifically amplify a fragment containing a tandemly repeated region. PCR was performed in 20 μl of a mixture containing 20 ng of genomic DNA, 67 mM Tris-HCl (pH 8.8), 16.6 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$, 6.7 μM ethylene-diamine tetraacetic acid, 10 mM β-mercaptoethanol, 1.5 mM deoxyribonucleoside triphosphates, 2 μl of dimethylsulfoxide, 10 pmol of each primer, and 1 unit of Ex-*Taq* DNA polymerase (Takara, Tokyo, Japan) using a set of primers (forward: 5′-ACCAGGAGG CCACCCCTG-3′ and reverse: 5′-TCCTGCAGCTTAGC CTTGGG-3′). Cycle conditions were as follows: 94°C for 2 min, then 35 cycles of 94°C for 30 s, 66°C for 30 s, and 72°C

Table 1. Allele frequencies estimated from genomic DNA of 46 unrelated Japanese and 31 Caucasian individuals

Allele	Size (bp)	Repeat (times)	Frequency	
			Japanese	Caucasian
a1	770	18	0.011	0.048
a2	737	17	0.109	0.016
a3	704	16	0.674	0.582
a4	671	15	0.065	0.145
a5	638	14	0.141	0.177
a6	605	13	0.000	0.032

for 1 min, with a final extension step of 4 min at 72°C, in a GeneAmp PCR 9700 System (Applied Biosystems, Foster City, CA, USA). The PCR products were electrophoresed on a 3% NuSieve 3:1 agarose gel (Takara).

Direct sequencing. Products obtained from the PCR experiments were used as templates for direct sequencing by the fluorescent dye-terminator cycle sequencing method (ABI PRISM Big Dye Terminator Cycle Sequencing Ready Reaction Kit; Perkin Elmer, Foster City, CA, USA).

Results and discussion

A tandemly repeated DNA sequence in exon 11 of the *CEL* gene corresponding to an allele containing 16 repeats is shown in Fig. 1. To investigate the polymorphic features of this segment, we performed PCR amplification and direct sequencing for unrelated Japanese and Caucasian individuals using a set of primers indicated in Fig. 1. A representative result of agarose-gel electrophoresis of PCR fragments shown in Fig. 2 clearly demonstrated variable numbers of this tandem repeat sequence (VNTR). Allelic frequencies of this VNTR in 46 unrelated Japanese and 31 unrelated Caucasian individuals were estimated from the size of the PCR fragments (Table 1). The observed frequency of heterozygotes was 44% in the Japanese population and 65% in the Caucasian population. We observed variations of 14–18 repeat units in the Japanese population and 13–18 units in the Caucasian population. The 13-repeat allele (designated as a6) was not observed in the 46 Japanese individuals examined. The 16-repeat allele for which we designated an a3 allele was the most frequent allele in the Japanese population (allelic frequency of 0.674) and in the Caucasian population (0.582). There seemed to be some differences in the allelic frequencies in the two populations, but the difference was not statistically significant.

During its transport from the endoplasmic reticulum to the *trans*-Golgi network, CEL is associated with intracellular membranes by means of a multiprotein folding complex. The association of CEL with membranes is required for the complete *O*-glycosylation of the protein (Nganga et al. 2000). Since this repeated sequence contains a potential site for *O*-linked glycosylation, which is required for CEL secretion (Bruneau et al. 1997), the difference in the number of

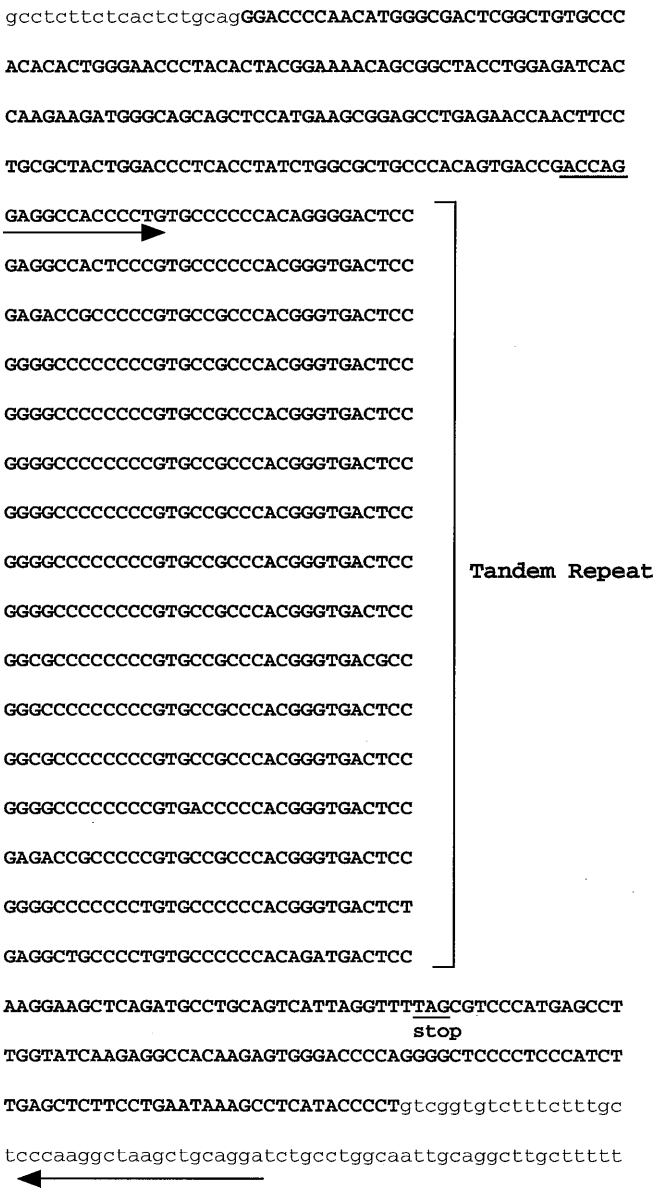

Fig. 1. Nucleotide sequences of the tandemly repeated region of exon 11 and its flanking introns in the *CEL* gene. The *uppercase letters* refer to exon sequences, and the *lowercase letters* to intronic sequences. Primer sequences for polymerase chain reaction (PCR) amplification used in this study are indicated by *arrows*. The translation termination codon is shown as *stop*

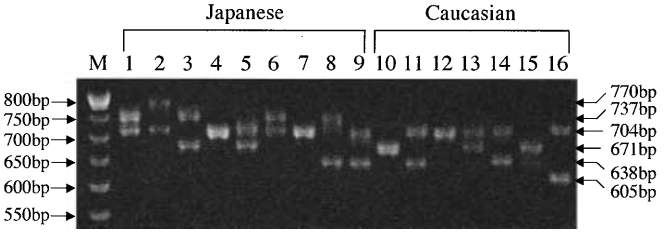

Fig. 2. Result of agarose-gel electrophoresis of PCR products corresponding to the tandemly repeated sequence. PCR products of nine Japanese and seven Caucasian individuals were analyzed. The result clearly indicates a variable number of tandem repeats type of polymorphism that reveals six alleles ranging from 770 bp to 605 bp. *M*, DNA size marker

repeated units may influence the function of the gene product. Moreover, this tandem-repeated fragment includes a mucin-like structure (Pasqualini et al. 1998). The mucin genes contain long stretches of VNTR polymorphisms in their coding region and the rare VNTR alleles of *MUC3* and *MUC7* were shown to be associated with the risk of ulcerative colitis (Kyo et al. 1999) and asthma (Kirkbride et al. 2001). Hence, the VNTR polymorphism in the *CEL* gene may be a useful genetic marker to study diseases associated with cholesterol absorption.

References

Bruneau N, Nganga A, Fisher EA, Lombardo D (1997) *O*-Glycosylation of C-terminal tandem-repeated sequences regulates the secretion of rat pancreatic bile salt-dependent lipase. J Biol Chem 272:27353–27361

Kirkbride HJ, Bolscher JG, Nazmi K, Vinall LE, Nash MW, Moss FM, Mitchell DM, Swallow DM (2001) Genetic polymorphism of *MUC7*: allele frequencies and association with asthma. Eur J Hum Genet 9:347–354

Kyo K, Parkes M, Takei Y, Nishimori H, Vyas P, Satsangi J, Simmons J, Nagawa H, Baba S, Jewell D, Muto T, Lathrop GM, Nakamura Y (1999) Association of ulcerative colitis with rare alleles of the human intestinal mucin gene, *MUC3*. Hum Mol Genet 8:307–311

Lombardo D, Guy O, Figarella C (1978) Purification and characterization of a carboxyl ester hydrolase from human pancreatic juice. Biochim Biophys Acta 527:142–149

Lombardo D, Fauvel J, Guy O (1980) Studies on the substrate specificity of a carboxyl ester hydrolase from human pancreatic juice. I. Action on carboxyl esters, glycerides and phospholipids. Biochim Biophys Acta 611:136–146

Mas E, Abouakil N, Roudani S, Franc JL, Montreuil J, Lombardo D (1993) Variation of the glycosylation of human pancreatic bile-salt-dependent lipase. Eur J Biochem 216:807–812

Nganga A, Bruneau N, Sbarra V, Lombardo D, Le Petit-Thevenin J (2000) Control of pancreatic bile-salt-dependent-lipase secretion by the glucose-regulated protein of 94 kDa (Grp94). Biochem J 352:865–874

Nilsson J, Bläckberg L, Carlsson P, Enerbäck S, Hernell O, Bjursell G (1990) cDNA cloning of human-milk bile-salt-stimulated lipase and evidence for its identity to pancreatic carboxylic ester hydrolase. Eur J Biochem 192:543–550

Pasqualini E, Caillol N, Panicot L, Mas E, Lloubes R, Lombardo D (1998) Molecular cloning of the oncofetal isoform of the human pancreatic bile salt-dependent lipase. J Biol Chem 273:28208–28218

Rogers S, Wells R, Rechsteiner M (1986) Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science 234:364–368

Taylor AK, Zambaux JL, Klisak I, Mohandas T, Sparkes RS, Schotz MC, Lusis AJ (1991) Carboxyl ester lipase: a highly polymorphic locus on human chromosome 9qter. Genomics 10:425–431