J Hum Genet (2002) 47:478–483

ORIGINAL ARTICLE

Yukio Watanabe · Akira Kinoshita · Takahiro Yamada
Tohru Ohta · Tatsuya Kishino · Naomichi Matsumoto
Mutsuo Ishikawa · Norio Niikawa · Koh-ichiro Yoshiura

A catalog of 106 single-nucleotide polymorphisms (SNPs) and 11 other types of variations in genes for transforming growth factor-β1 (TGF-β1) and its signaling pathway

Received: May 9, 2002 / Accepted: June 4, 2002

Abstract Transforming growth factor-β1 (TGF-β1) is a multifunctional cytokine that is produced in the platelet, bone, placenta, and other tissues. It acts as a growth inhibitor in many types of cells, and also mediates extracellular matrix production and immunosuppression. Mutations in the specific domain of its gene (*TGFB1*) cause Camurati-Engelmann disease, a bone-sclerosing disorder, and those in other domains may be associated with osteoporosis. We identified 106 single-nucleotide polymorphisms and 11 other types of variations in *TGFB1* and six other genes. These genes were TGF-β type I receptor gene *(TGFBR1*), TGF-β type II receptor gene *(TGFBR2),* SMAD2 gene *(SMAD2),* SMAD3 gene *(SMAD3),* SMAD4 gene *(SMAD4),* and SMAD7 gene *(SMAD7)*, all of which compose the TGF-β1 signaling pathway. We also estimated allele frequencies of these DNA polymorphisms among 48 Japanese individuals. Our data will provide a useful resource for the study of disease susceptibility.

Key words Single-nucleotide polymorphism (SNP) · TGF-β1 · TGF-β1 signaling pathway · TβRs · SMADs Japanese population · Allele frequency

Y. Watanabe · A. Kinoshita · T. Yamada · N. Matsumoto · N. Niikawa · K-i. Yoshiura (✉)
Department of Human Genetics, Nagasaki University School of Medicine, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan
Tel. +81-95-849-7120; Fax: +81-95-849-7121
e-mail: kyoshi@net.nagasaki-u.ac.jp

Y. Watanabe · M. Ishikawa
Department of Obstetrics and Gynecology, Asahikawa Medical College, Asahikawa, Japan

T. Yamada
Department of Obstetrics and Gynecology, Hokkaido University School of Medicine, Sapporo, Japan

T. Ohta · T. Kishino
Gene Research Center, Nagasaki University, Nagasaki, Japan

Y. Watanabe · A. Kinoshita · T. Yamada · T. Ohta · T. Kishino · N. Matsumoto · N. Niikawa · K-i. Yoshiura
CREST, Japan Science and Technology Corporation, Kawaguchi, Japan

Introduction

Transforming growth factor-β1 (TGF-β1) is a multifunctional cytokine that is essential for maintaining homeostasis involving bone and muscle differentiation, immune response, and tumor suppression (Massagué 1998; Alliston et al. 2001; Centrella et al. 1994; Brennan et al. 1991; Liu et al. 2001; Werner et al. 2000; Derynck et al. 2001). The role of TGF-β1 in bone formation was confirmed by our previous work that showed that mutations in the latency-associated polypeptide domain of its gene (*TGFB1*) cause Camurati-Engelmann disease, a heritable bone-sclerosing disorder (Kinoshita et al. 2000; Saito et al. 2001). In addition, an inverse-looking condition, osteoporosis, may be associated with *TGFB1*. Yamada et al. (1999) demonstrated in Japanese osteoporotic patients such an association in an exonic single-nucleotide polymorphism (SNP), namely, a C→T substitution at nucleotide (nt) position 29 in *TGFB1*. A similar association was obtained in an Italian population, although a polymorphism (713-8delC) in the gene showing the association was different (Bertoldo et al. 2000).

At least six molecules other than TGF-β1 have been known to participate in the TGF-β1 signaling pathway. They include TGF-β type I receptor (TβR-I), TGF-β type II receptor (TβR-II), SMAD2, SMAD3, SMAD4, and SMAD7. Thus, identification of nucleotide polymorphisms in genes *TGFB1, TGFBR1, TGFBR2, SMAD2, SMAD3, SMAD4,* and *SMAD7* for these seven proteins may be useful to clarify osteoporotic disorders.

In this article, we provide a series of fine-scale maps of variations in these genes. These maps contain a total of 106 SNPs and 11 other types of variants in a Japanese population.

Subjects and methods

Seven genes (*TGFB1, TGFBR1, TGFBR2, SMAD2, SMAD3, SMAD4,* and *SMAD7*) were studied for their

genetic polymorphisms. The genes are those for molecules composing the TGFβ-1 signaling pathway. Blood samples were obtained with informed consent from 48 healthy Japanese individuals in Nagasaki prefecture. On the basis of genomic sequences from GenBank database, we designed primers to amplify all exons of these genes as well as sequences at 1.5-kb 5′- and 3′-flanking regions and 5′- and 3′-untranslated regions (UTRs). Genomic DNA was prepared from lymphocytes by standard phenol/chloroform extraction and isopropanol precipitation, and 10 ng of it was used as a template for polymerase chain reaction (PCR).

PCR was performed first using DNA from 16 normal individuals in a 10-μl reaction volume, containing 500 nM of each primer, 0.2 units of Takara ExTaq HS-version (Takara, Otsu, Japan), ExTaq buffer, and a 200-μM deoxyribonucleosido triphosphate mixture. PCR products were purified using a PCR purification kit (Qiagen, Hilden, Germany) or ExoSAP-IT (USB, Cleveland, OH, USA) according to the manufacturers' protocols. The purified PCR products were sequenced using the Big-Dye terminator kit version 3.0 (PE Applied Biosystems, Foster City, CA, USA) and run on an Autosequencer (ABI PRISM 3100, PE Applied Biosystems). The electropherogram was aligned by Autoassembler software (PE Applied Biosystems), and SNPs were detected on the aligned electropherogram.

Sequence analysis was extended to a total of 48 persons (96 chromosomes), and allele frequency was then calculated.

Results

We identified a total of 106 SNPs and 11 other types of variations in the seven genes among 48 healthy Japanese individuals. Fine physical maps of these genes are shown in Fig. 1, and detailed information on the genetic variations is summarized in Tables 1 and 2. Of the 106 SNPs identified, 79 (75%) had not been deposited in the SNP database (dbSNP) of the U.S. National Center for Biotechnology Information, and 9 (TGFB1-6, TGFB1-7, TGFB1-8, TGFB1-10, TGFBR1-6, SMAD3-4, SMAD3-16, SMAD3-17, and SMAD3-18) were observed within regions that correspond to repetitive sequences predicted by the RepeatMasker Program (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker). Five SNPs (TGFBR2-4, SMAD4-7, SMAD7-8, SMAD7-9, SMAD7-14) and one insertion (SMAD7-19) were rare polymorphisms that were not identified in 16 persons at first, but they were unexpectedly identified near the common SNPs during the sequencing for 48 persons. Of five SNPs identified within coding regions of the genes, two were nonsynonymous and predicted to result in amino-acid substitutions that might affect the structure and biological function of the respective gene products. Detailed information of the 106 SNPs and 11 other variants in the seven genes is provided in the following paragraphs.

SNPs in *TGFB1*

A total of 11 SNPs were identified in *TGFB1*. They included one at the 5′-flanking region, one in the coding region and nine in introns. The only SNP (TGFB1-2, 29C→T) found in the coding region (exon 1) was nonsynonymous and predicted to change from proline to leucine (P10L). This change has already been observed in another Japanese population (Yamada et al. 2001). Frequencies of C and T alleles at the TGFB1-2 locus among the 48 individuals were estimated to be 58.3% and 41.7%, respectively. A deletion polymorphism (one-base deletion in intron 4, 713-8delC), frequently observed in an Italian population (Bertoldo et al. 2000), was not found in our study.

Polymorphisms in *TGFBR1* and *TGFBR2*

Thirteen SNPs were identified in *TGFBR1,* including 12 in introns, and 1 in the 3′-UTR. In *TGFBR2,* 31 SNPs were identified: 1 in exon 4, 27 in introns, and 3 in the 3′-UTR. The SNP (TGFBR2-18, 571G→A) in exon 4 was nonsynonymous (V191I), although the A allele was very rare with a frequency of 1.0%. Two deletion polymorphisms were also identified in the *TGFBR2* genomic region.

Polymorphisms in the *SMAD* gene family

All six SNPs identified in *SMAD2* were located in its introns. In *SMAD3*, a total of 21 SNPs were identified. They included 1 (SMAD3-5) in exon 2, 1 each in the 5′-flanking and in the 3′-UTRs, and 18 SNPs in introns. The SMAD3-5 was synonymous (A309G, L103L). Other variations observed in *SMAD3* were one deletion (SMAD3-22), 2 insertions (SMAD3-23, and SMAD3-24), and one tandem repeat polymorphism (SMAD3-25). In *SMAD4*, there were two SNPs in the 5′-flanking region, four in introns, and two in the 3′-UTR. In addition, one deletion polymorphism (SMAD4-9) was identified in its intron 3. In *SMAD7*, a total of 16 SNPs were identified, 2 (SMAD7-11, and SMAD7-12) in the coding region, 10 in introns, one in the 3′-UTR, and 3 in the 3′-flanking region. Both SMAD7-11 and SMAD7-12 were synonymous (L298L and G402G, respectively). Two deletions (SMAD7-17 and SMAD7-18) and two insertions (SMAD7-19 and SMAD7-20) were also identified.

Discussion

We have identified a total of 106 SNPs and 11 deletion/insertion/tandem-repeat polymorphisms in the seven genes that participate in the TGF-β1 signaling pathway. Because the pathway concerns bone and muscle differentiation, immunosuppression, and carcinogenesis, the variants we identified are useful for such research fields. It has repeatedly been reported that alterations of the pathway affect various human disorders. A lack of growth inhibition by TGF-β leads to carcinogenesis (Massagué 1998), and excess TGF-β activity may play a significant role in the pathogen-

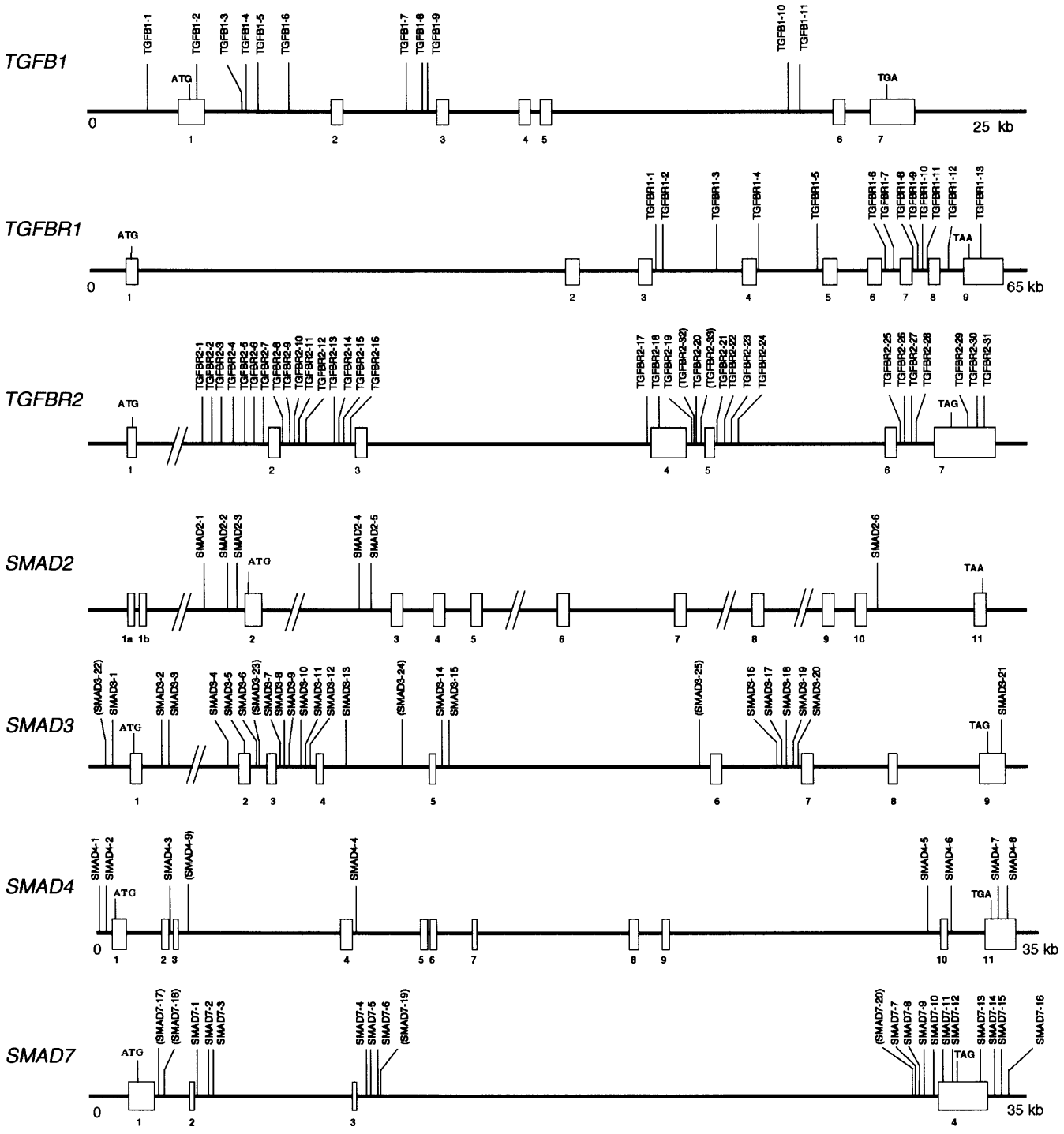

Fig. 1. Genomic organizations of seven genes composing the TGF-β1 signaling pathway, and locations of Single-nucleotide polymorphisms (SNPs) in the genes: the TGF-β1 gene (*TGFB1*), the transforming growth factor-β type I receptor gene (*TGFBR1*), the transforming growth factor-β type II receptor gene (*TGFBR2*), the SMAD2 gene (*SMAD2*), the SMAD3 gene (*SMAD3*), the SMAD4 gene (*SMAD4*), and the SMAD7 gene (*SMAD7*). Exons (their numbers are *below lines*) are depicted by *open boxes*, and SNPs and other variants (in *parentheses*) are indicated *above lines*. Detailed information for the polymorphisms in each gene is shown in Tables 1 and 2. Total length of *TGFBR2, SMAD2*, and *SMAD3* was unknown because of gaps in DNA sequences

esis of fibrotic disorders in the kidney (Border et al. 1992) and in inflammatory disorders (Shull et al. 1992; Kulkarni et al. 1993). As mentioned earlier, germline mutations in *TGFB1* cause a heritable bone sclerosing disorder, Camurati-Engelmann disease (Kinoshita et al. 2000; Saito et al. 2001). Moreover, certain polymorphisms in *TGFB1* have recently been reported to be associated with the degree of bone mineral density and susceptibility to osteoporosis (Bertoldo et al. 2000; Yamada et al. 2001), as well as with asthma severity (Pulleyn et al. 2001). Somatic

Table 1. SNPs and other polymorphisms in *TGFB1*, *TGFBR1*, and *TGFBR2* in a Japanese population

Gene and symbol of variants	Position (nt)[a]	Base changes (large case), and their 5' and 3' flanking sequences (small case)	Amino acid substitution	Allele frequency (allele)	Accession number in GenBank	dbSNP in NCBI	Repetitive sequence
TGFB1					AC011462.4		
TGFB1-1	5'-FR (-1347)	gcctcctgacccttccatcc **T/C** tcaggtgtcctgttgccccc		0.417 (C)	76571	rs1800469	
TGFB1-2	exon 1 (29)	ctccgggctgcggctgctgc **C/T** gctgctgctaccgctgctgt	P10L	0.417 (T)	77946	rs1982073	
TGFB1-3	intron1 (1511)	agcgtccggcggcctcatcc **C/T** cctcccttccctttccttcc		0.427 (T)	79428		
TGFB1-4	intron1 (1546)	ccttcccatgccccccggcgg **A/G** ggcggggatcgctcgcggag		0.427 (G)	79463		
TGFB1-5	intron1 (2064)	atttttctcctccacggtcc **T/G** gttgcctcgtctccgtctct		0.438 (G)	79981	rs2241715	
TGFB1-6	intron1 (3435)	cgtccggcaccatgcctggc **C/T** aatttttgtatttttagtag		0.396 (T)	81352	rs284658	Alu
TGFB1-7	intron2 (2085)	ttacaggcatgcaccaccac **A/G** cctagctaattttgtatttt		0.438 (G)	84752	rs2014015	Alu
TGFB1-8	intron2 (2484)	gtcttgggttcaagtgattc **T/C** cctgcctcagcctcctgagt		0.344 (C)	85145		Alu
TGFB1-9	intron2 (2691)	cagccggaatcattagaaat **G/T** acttctaagttactgagaat		0.406 (T)	85358		
TGFB1-10	intron5 (7219)	tgcctcagactccccagtag **C/G** tgggattacaggtgcccgcc		0.042 (G)	96298		Alu
TGFB1-11	intron5 (8157)	tagtacctactgcattccag **A/G** cactgctttaggagttaagg		0.177 (G)	97236		
TGFBR1					AL162427.24		
TGFBR1-1	intron3 (363)	aagatattcacatgatctgc **A/G** gctaacccattctattcagg		0.010 (G)	60238		
TGFBR1-2	intron3 (461)	aattgtgaaggttatttact **T/A** tttaaacattatatgaaatt		0.313 (A)	60140		
TGFBR1-3	intron3 (3737)	gtgtggccctctacaaggct **T/G** tcataaaatgtaataaatta		0.458 (C)	56864		
TGFBR1-4	intron4 (576)	tgctggtgcctgaggcccag **T/C** gtcaggtcagatgctgaaga		0.469 (C)	54675		
TGFBR1-5	intron4 (3569)	cgtccatcttttactggatt **G/A** tgctgtctcctaatgttgta		0.438 (A)	51682		
TGFBR1-6	intron6 (254)	taataacagtagtaactaac **G/A** tttatagggcacctactatg		0.115 (A)	48198		MIR
TGFBR1-7	intron6 (1195)	cagtatgaggactggcattc **T/G** tttgtatctattatttttt		0.500 (G)	47257	rs334353	
TGFBR1-8	intron7 (24)	aattgctctcctctccccca **A/G** tagtttgtcatgagcagaag		0.458 (G)	46707	rs334354	
TGFBR1-9	intron7 (395)	tctccaacttgatagttaca **A/G** ttacatacacagaatacaat		0.490 (G)	46336	rs334355	
TGFBR1-10	intron7 (481)	ctagctttttagatggagta **C/T** ttttaatgtttactcaggaa		0.042 (T)	46250	rs334356	
TGFBR1-11	intron7 (766)	cacaatagggttctagggga **A/G** aatgtcaatctggtcatcac		0.458 (G)	45965	rs1888225	
TGFBR1-12	intron8 (547)	tttccagcagtatctgagct **G/T** ctcaacttcagctctaacaa		0.052 (T)	45009	rs334358	
TGFBR1-13	3'-UTR (69)	ggttttaatttgggaggtca **A/G** ttgttctacctcactgagag		0.042 (G)	43966	rs868	
TGFBR2					AC096921.2		
TGFBR2-1	intron1 (-1399)	agtattaattcatgtttgct **G/C** gttgagtgcctagtagcctt		0.031 (C)	117969		
TGFBR2-2	intron1 (-1332)	catgtcttaaacatcatcac **T/C** atctcattgaggaagaaaat		0.344 (C)	118036		
TGFBR2-3	intron1 (-1139)	tgtatcctatcctacatggc **A/T** tcaagggccaaacatggact		0.021 (T)	118229		
TGFBR2-4	intron1 (-1137)	tatcctatcctacatggcat **C/T** aagggccaaacatggactcc		0.010 (T)	118231		
TGFBR2-5	intron1 (-1111)	ccaaacatggactccttgta **T/G** tcatcctgtgttcactcatt		0.031 (G)	118257		
TGFBR2-6	intron1 (-801)	aacttggcttcttcccagaa **G/A** ggggcaggagattggtgatg		0.031 (A)	118567		
TGFBR2-7	intron1 (-619)	tgcctcaaacacagagagag **A/G** tcagtagaactagtcccttg		0.031 (G)	118749		
TGFBR2-8	intron2 (7)	tgtgtggctgtatggtaagc **G/A** agccttttaagaagttattc		0.281 (A)	119543		
TGFBR2-9	intron2 (220)	gtatataaacgtatcatgat **C/T** ataaaatgatcatgataaat		0.281 (T)	119756		
TGFBR2-10	intron2 (240)	tataaaatgatcatgataaa **C/T** gtaaagatcataaaatgatc		0.281 (T)	119776		
TGFBR2-11	intron2 (333)	cgcttagtagtcatttaacc **A/G** acacctcctcagcaccccat		0.281 (G)	119869		
TGFBR2-12	intron2 (375)	gtgagaccaggagctttgtt **C/G** tcttgaggcaatctcaggtc		0.031 (G)	119911		
TGFBR2-13	intron2 (4251)	ccagattagcccttcctcc **G/T** ctccctctctagacagagag		0.406 (T)	123787		
TGFBR2-14	intron2 (4341)	aaagggtggatgtcagacta **G/A** ccttgttggacagtcatcca		0.406 (A)	123877		
TGFBR2-15	intron2 (4903)	aggcatgctctttgaaaaca **G/A** tatagctagacaggacagat		0.406 (A)	124439		
TGFBR2-16	intron2 (4912)	ctttgaaaacagtatagcta **G/A** acaggacagattttgccact		0.406 (A)	124448		
TGFBR2-17	intron3 (21274)	ctccttgttttgtttcccca **T/A** cagaatataacaccagcaat		0.406 (A)	146255		
TGFBR2-18	exon4 (571)	tcatcttctactgctaccgc **G/A** ttaaccggcagcagaagctg	V191I	0.010 (A)	146375		
TGFBR2-19	intron4 (657)	cagtggagtccatggagagc **A/C** gcactgtcttctttagccag		0.375 (C)	147715		
TGFBR2-20	intron4 (1226)	acaaagattaagactcagtc **C/T** taaagtgatcaaaactgata		0.396 (T)	148284		
TGFBR2-21	intron5 (239)	tataaacctttgtataatta **C/T** tcttctcttgacacatctct		0.125 (T)	149106		
TGFBR2-22	intron5 (521)	tctagtgaagaggttattga **G/A** gaaggccgtcttcttactgc		0.125 (A)	149388		
TGFBR2-23	intron5 (550)	tcttcttactgcttaagcct **G/C** caagactgattaggagacct		0.292 (C)	149417		
TGFBR2-24	intron5 (864)	agaaggcaagaacagttgag **G/A** tctacattactataccctt		0.385 (A)	149731		
TGFBR2-25	intron6 (236)	agccattcggaggacatttg **C/T** cacattttccaagcccccctt		0.396 (T)	163368	rs2276768	
TGFBR2-26	intron6 (686)	ccatagcagggcagtcattc **T/C** ggactcataaagaggctgat		0.031 (C)	163818		
TGFBR2-27	intron6 (869)	tggccctatctcactgagaa **A/G** actgttctttctactgacta		0.469 (G)	164001		
TGFBR2-28	intron6 (1120)	ggttggcgaggggggaaccag **G/C** ctgtggtgggtcacaggctt		0.125 (C)	164252		
TGFBR2-29	3'-UTR (747)	gggctagtttagaaaactctc **C/G** ctcaacctagtttagaaact		0.031 (G)	166967		
TGFBR2-30	3'-UTR (889)	ccatctactaatgaaaaatt **G/C** ttctttttttcatctttccc		0.021 (C)	167109		
TGFBR2-31	3'-UTR (988)	attctcactctaggctttat **C/T** gtgtttacttttttcattaca		0.010 (T)	167208		
TGFBR2-32	intron4 (1141 - 5)	aaaatgcttttatacaattt **DEL*** agtagagatgttatgcaatt		0.417 (DEL)	148199-203		
TGFBR2-33	intron4 (1296)	agtttcaggaataaaaaaaa **DEL(A)** ttcaattttggaaagcaaaa		0.115 (DEL)	148354		

SNPs, Single-nucleotide polymorphisms; nt, nucleotide; dbSNP, database of SNPs; NCBI, National Center for Biotechnology Information; FR, flanking region; UTR, untranslated region; DEL, deletion; DEL*, del(AATTT)

[a] Nucleotide numbering is according to the mutation nomenclature (den Dunnen and Antonarakis 2000)

Table 2. SNPs and other polymorphisms in *SMAD2*, *SMAD3*, *SMAD4*, and *SMAD7* in a Japanese population

Gene and symbol of variants	Position (nt)[a]	Base changes (large case), and their 5' and 3' flanking sequences (small case)	Amino acid substitution	Allele frequency (allele)	Accession number in GenBank	dbSNP in NCBI	Repetitive sequence
SMAD2					AP001386		
SMAD2-1	intron1 (-1024)	gaaggcaaaggaaagaaaag **T/C** gtaagagctaagatattaag		0.427 (C)	63198	rs1942158	
SMAD2-2	intron1 (-440)	ttcttgtaactgtagatttt **G/C** cctcctattagttatgaaag		0.302 (C)	62614		
SMAD2-3	intron1 (-147)	attcatttaaattcataaag **A/G** taacgttttcatgggtggag		0.271 (G)	62321		
SMAD2-4	intron2 (-688)	aaagttccattccaaaactt **C/G** cttttaagctctctttccct		0.302 (G)	104892		
SMAD2-5	intron2 (-622)	gaaatgaaataaaggtaggg **A/G** aaaagggagaattatctctg		0.302 (G)	104958		
SMAD2-6	intron10 (202)	tatggaagatattaaaacat **A/G** cacaggcgtgcatttttaacg		0.406 (G)	44985	rs1787186	
SMAD3					AC087482.4		
					AC012568.7		
SMAD3-1	5'-FR (-1719)	ttttttttaacaaagcaggg **G/A** tgggggtgggagattcctgc		0.240 (A)	158643[b]		
SMAD3-2	intron1 (1040)	tgcgggtaccttggcttcct **A/G** tctctacctctcagcctcct		0.333 (G)	161607[b]		
SMAD3-3	intron1 (1185)	aaagggattgtcatccctgc **C/T** cctgggccagggtggactcc		0.333 (T)	161752[b]		
SMAD3-4	intron1 (-967)	cagcctgagtgacagagcaa **G/C** accctgtatcaaaagtaaaa		0.187 (C)	79183	rs1197674	Alu
SMAD3-5	exon2 (309)	ctgcacagccaccacgagct **G/A** cgggccatggagctgtgtga	L103L	0.187 (A)	80252	rs1065080	
SMAD3-6	intron2 (59)	aggggtcatcacctctcccc **G/C** gctccccatccccccgaggg		0.417 (C)	80402	rs2289261	
SMAD3-7	intron3 (85)	cctgtggccccaatctctgc **C/T** ccctggccgtccccccgctca		0.083 (T)	80724	rs2289260	
SMAD3-8	intron3 (128)	ccctctttgcgcacagctct **G/A** gcctgagggccctgactca		0.302 (A)	80767	rs2289259	
SMAD3-9	intron3 (430)	tcgcctgggcatcaggcctc **G/A** gtgagggggctccaacctggg		0.406 (A)	81069		
SMAD3-10	intron3 (1083)	tctggacgcctccctggagg **G/A** ggtggggcttaaccctcacc		0.427 (A)	81722		
SMAD3-11	intron3 (1208)	aaaggcagttatgatccaag **T/C** gggagtcagaggtggacagg		0.187 (C)	81847		
SMAD3-12	intron3 (1301)	acctggagctcctacagcca **C/T** ggatgctagcatcatggtgt		0.313 (T)	81930		
SMAD3-13	intron4 (818)	gtgatgtgtgtgctcctgcc **G/A** tgaggggcagggcttatttc		0.375 (A)	82926		
SMAD3-14	intron5 (987)	tagttaatatttcttggcga **A/G** tcaaagtggattggactggt		0.021 (G)	86846		
SMAD3-15	intron5 (1071)	tgaagaaactcatcatttgg **A/G** atattaggagatgcttgaaa		0.260 (G)	86930		
SMAD3-16	intron6 (2944)	gattattggtgtgagccacc **G/A** cgcctggccccatccctat		0.490 (A)	99652	rs2955780	Alu
SMAD3-17	intron6 (2965)	cgcctggccccatccctat **C/T** gtagagacaaacgaagactc		0.490 (T)	99673		MIR
SMAD3-18	intron6 (2977)	atccctatcgtagagacaaa **C/T** gaagactcagagggtgaata		0.490 (T)	99685		MIR
SMAD3-19	intron6 (3161)	ccgtttgcctggggaagctg **G/T** cagtcactgggagcagctct		0.490 (T)	99869	rs2289791	
SMAD3-20	intron6 (3179)	tggcagtcactgggagcagc **T/C** ctgctgttctgcctccttg		0.490 (C)	99887	rs2289791	
SMAD3-21	3'-UTR (402)	tggagttcaccttggaaggg **C/T** gttctaggtaggaagagccc		0.490 (T)	106193		
SMAD3-22	5'-FR (-1732)	tccagataactttttttttt **DEL(T)** aacaaagcagggtgggggt		0.177 (DEL)	158630[b]		
SMAD3-23	intron2 (59-60)	ggggtcatcacctctccccg **INS(C)** gctccccatccccccgaggg		0.021 (INS)	80402-3		
SMAD3-24	intron4 (3021-2)	accatgtgagttctttttttt **INS(T)** ctggaggtgagaggtgtaga		0.365 (INS)	85129-30		
SMAD3-25	intron5 (10369)	ggggcagggcctttatgagc **(TAAA)n** aagagaaatcaatggccctt		0.146 (TAAA)8	96228		
				0.375 (TAAA)9			
				0.396 (TAAA)10			
				0.052 (TAAA)11			
				0.010 (TAAA)12			
SMAD4					AC091551.11		
SMAD4-1	5'-FR (-433)	gctggtaagattttccttta **G/A** ggtgactagtaaagtcagta		0.427 (A)	94865		
SMAD4-2	5'-FR (-262)	aagaattttaagtaattttc **A/C** actctgagcatcaaatttta		0.042 (C)	95036		
SMAD4-3	intron2 (159)	tatacgtatttaaatttgaa **T/C** ttaggaacaaacttatattt		0.490 (C)	97271	rs2276163	
SMAD4-4	intron4 (238)	tactttactattacatgtag **C/A** agtaagtagcctcttaagct		0.479 (A)	103486		
SMAD4-5	intron9 (8858)	gtcccatatacccatcgcac **G/A** gattaagtagttacaatttt		0.427 (A)	124308		
SMAD4-6	intron10 (257)	attattatattatactgttc **T/C** atctcctccttttttgttga		0.479 (C)	125300	rs2298617	
SMAD4-7	3'-UTR (79)	ctgtttataatctgaagata **T/C** atttcacttttgttctgctt		0.010 (C)	126813		
SMAD4-8	3'-UTR (412)	ttctcaaagttaattcacct **A/G** tgttattttgtgtacaagtt		0.031 (G)	127146		
SMAD4-9	intron3 (242)	aactaaagtactgtaaaaaa **DEL(A)** tatgatagaggaatgcacca		0.458 (DEL)	97818		
SMAD7					AC024384.4		
SMAD7-1	intron2 (8)	ctcaaaccaactggtgagta **G/A** atgatttttaaaatatcctg		0.281 (A)	163602		
SMAD7-2	intron2 (562)	tctcagtattcgccatagggg **A/G** ctatcaaagtttggagcata		0.437 (G)	163048		
SMAD7-3	intron2 (576)	atagggggctatcaaagtttg **A/G** agcatagctctgcctgtctg		0.365 (G)	163034		
SMAD7-4	intron3 (493)	caagttctgattcccagccc **A/G** gcctgctgcaggattttgca		0.188 (G)	157213		
SMAD7-5	intron3 (496)	gttctgattcccagcccagc **C/T** tgctgcaggattttgcagca		0.323 (T)	157210	rs1873190	
SMAD7-6	intron3 (601)	tagcggggcaggttttcctc **A/G** gggctcacaggtatgcgtga		0.313 (G)	157105	rs1873191	
SMAD7-7	intron3 (20004)	tttatttttaccagataggt **C/T** cctcatgaccacggcactga		0.021 (T)	137702		
SMAD7-8	intron3 (20060)	acaggcacccagagaccatc **C/G** cataggtggtcccagagtct		0.010 (G)	137646		
SMAD7-9	intron3 (20074)	accatcgcataggtggtccc **G/A** gagtctagacacctgggttc		0.010 (A)	137632		
SMAD7-10	intron3 (20493)	ctccttctttctgtctcgtc **C/T** tgagagatgagctggtgact		0.135 (T)	137213		
SMAD7-11	exon4 (894)	caggggaatggcttttgcct **C/T** ggacagctcaattcggacaa	L298L	0.135 (T)	136956		
SMAD7-12	exon4 (1206)	cagatcagctttgtgaaggg **C/G** tggggccagtgctacacccg	G402G	0.135 (G)	136644		
SMAD7-13	3'-UTR (1153)	ttctaactacaaaggtttaa **A/G** tgaacaagagaagcattctc		0.167 (G)	135416		
SMAD7-14	3'-FR (2053)	ggactaggccacctcctggc **G/A** ccctcccaaatactgtcctg		0.010 (A)	134516		
SMAD7-15	3'-FR (2116)	aagttccactctgcaatatc **G/C** tccttgtctgctgggactt		0.063 (C)	134453		
SMAD7-16	3'-FR (2247)	gccttggttcgctgctgaac **C/T** gtgatctgagccatgtttta		0.063 (T)	134322		
SMAD7-17	intron1 (143-64)	gtgtgcataccctcctggga **DEL*** agagactgggggctgaaagt		0.010 (DEL)	164874-95		
SMAD7-18	intron1 (177)	tcctgggaagagactggggg **DEL(C)** tgaaagtagagaagggacaa		0.010 (DEL)	164861		
SMAD7-19	intron3 (618-9)	tcagggctcacaggtatgcg **INS**** tgaaggcctggaatgccagt		0.021 (INS)	157102-3		
SMAD7-20	intron3 (20001-2)	tttttatttttaccagatag **DEL***** gtccctcatgaccacggcac		0.125 (DEL)	137706-7		

SNPs, Single-nucleotide polymorphisms; nt, nucleotide; dbSNP, database of SNPs; NCBI, National Center for Biotechnology Information; FR, flanking region; UTR, untranslated region; DEL, deletion; INS, insertion; DEL*, del(GAGTGTGCATACCCTCCTGGGA); INS**, ins(GCTCACAGGTATGCG); DEL***, del(AGAG)

[a] Nucleotide numbering is according to the mutation nomenclature (den Dunnen and Antonarakis 2000)

[b] 158643, 161607, 161752, 158630 are retrieved from AC087482.4

mutations in *TGFBR2* were found in gastrointestinal cancers with microsatellite instability (Myeroff et al. 1995). Inactivated *SMAD2* and *SMAD4* caused by nucleotide alterations, and regional or entire chromosomal deletions involving the genes, were observed to be associated with pancreatic cancer (Hahn et al. 1996). Originally, *SMAD4* was identified as a candidate tumor suppressor gene, and a large number of all human pancreatic cancers are caused by somatic deletions or mutations of *SMAD4*, also referred to as *DPC4* (the deleted in pancreatic carcinoma locus 4) (Hahn et al. 1996). In addition, abnormalities of TGF-β1 and other molecules composing the signaling pathway are involved in many other disorders (Myeroff et al. 1995; Knaus et al. 1996; Eppert et al. 1996; Schutte et al. 1996).

In conclusion, our collection of genetic variations for the seven genes provides useful genetic data for research, not only on the TGF-β1 signaling pathway, but also on its related disorders. Identification of genes that concern such disorders will contribute to future personalized medical service.

Acknowledgments We are grateful to all persons who participated in this study. We also thank Ms. Kazumi Miyazaki, Naoko Takaki, Yasuko Noguchi, and Naoko Yanai for their excellent technical assistance. This study was supported by CREST, Japan Science and Technology Corporation (JST), and in part by Grants-in-Aid for Scientific Research (Nos. 13203002, 13671787, 1367014, 12204140, and 13854024) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Alliston T, Choy L, Ducy P, Karsenty G, Derynck R (2001) TGF-β-induced repression of CBFA1 by Smad3 decreases cbfa1 and osteocalcin expression and inhibits osteoblast differentiation. EMBO J 20:2254–2272

Bertoldo F, D'Agruma L, Furlan F, Colapietro F, Lorenzi MT, Maiorano N, Iolascon A, Zelante L, Locascio V, Gasparini P (2000) Transforming growth factor-beta1 gene polymorphism, bone turnover, and bone mass in Italian postmenopausal women. J Bone Miner Res 15:634–639

Border WA, Ruoslahti E (1992) Transforming growth factor-β in disease: the dark side of tissue repair. J Clin Invest 90:1–7

Brennan TJ, Edmondson DG, Li L, Olson EN (1991) Transforming growth factor β represses the actions of myogenin through a mechanism independent of DNA binding. Proc Natl Acad Sci USA 88:3822–3826

Centrella M, Horowitz MC, Wozney JM, McCarthy TL (1994) Transforming growth factor-β gene family members and bone. Endocr Rev 15:27–39

den Dunnen JT, Antonarakis SE (2001) Nomenclature for the description of human variations. Hum Genet 109:121–124

Derynck R, Akhurst RJ, Balmain A (2001) TGF-β signaling in tumor suppression and cancer progression. Nat Genet 29:117–129

Eppert K, Scherer SW, Ozcelik H, Pirone R, Hoodless P, Kim H, Tsui LC, Bapat B, Gallinger S, Andrulis IL, Thomsen GH, Wrana JL, Attisano L (1996) MADR2 maps to 18q21 and encodes a TGFβ-regulated MAD-related protein that is functionally mutated in colorectal carcinoma. Cell 86:543–552

Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, Rozenblum E, Weinstein CL, Fischer A, Yeo CJ, Hruban RH, Kern SE (1996) DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 271:350–353

Kinoshita A, Saito T, Tomita H, Makita Y, Yoshida K, Ghadami M, Yamada K, Kondo S, Ikegawa S, Nishimura G, Fukushima Y, Nakagomi T, Saito H, Sugimoto T, Kamegaya M, Hisa K, Murray JC, Taniguchi N, Niikawa N, Yoshiura K (2000) Domain-specific mutations in TGFB1 result in Camurati-Engelmann disease. Nat Genet 26:19–20

Knaus PI, Lindemann D, DeCoteau JF, Perlman R, Yankelev H, Hille M, Kadin ME, Lodish HF (1996) A dominant inhibitory mutant of the type II transforming growth factor β receptor in the malignant progression of a cutaneous T-cell lymphoma. Mol Cell Biol 16:3480–3489

Kulkarni AB, Huh C, Becker D, Geiser A, Lyght M, Flanders KC, Roberts AB, Sporn MB, Ward JM, Karlsson S (1993) Transforming growth factor β1 null mutation in mice causes excessive inflammatory response and early death. Proc Natl Acad Sci USA 90:770–774

Liu D, Black BL, Derynck R (2001) TGF-β inhibits muscle differentiation through functional repression of myogenic transcription factors by Smad3. Genes Dev 15:2950–2966

Massagué J (1998) TGF-β signal transduction. Annu Rev Biochem 67:753–791

Myeroff LL, Parsons R, Kim SJ, Hedrick L, Cho KR, Orth K, Mathis M, Kinzler KW, Lutterbaugh J, Park K, Bang YJ, Lee HY, Park JG, Lynch HT, Roberts AB, Vogelstein B, Markowitz SD (1995) A transforming growth factor β receptor type II gene mutation common in colon and gastric but rare in endometrial cancers with microsatellite instability. Cancer Res 55:5545–5547

Pulleyn LJ, Newton R, Adcock IM, Barnes PJ (2001) TGFβ1 allele association with asthma severity. Hum Genet 109:623–627

Saito T, Kinoshita A, Yoshiura K, Makita Y, Wakui K, Honke K, Niikawa N, Taniguchi N (2001) Domain-specific mutations of a transforming growth factor (TGF)- β1 latency-associated peptide cause Camurati-Engelmann disease because of the formation of a constitutively active form of TGF-β1. J Biol Chem 276:11469–11472

Schutte M, Hruban RH, Hedrick L, Cho KR, Nadasdy GM, Weinstein CL, Bova GS, Isaacs WB, Cairns P, Nawroz H, Sidransky D, Casero RA Jr, Meltzer PS, Hahn SA, Kern SE (1996) *DPC4* gene in various tumor types. Cancer Res 56:2527–2530

Shull MM, Ormsby I, Kier AB, Pawlowski S, Diebold RJ, Yin M, Allen R, Sidman C, Proetzel G, Calvin D, Annunziata N, Doetschman T (1992) Targeted disruption of the mouse transforming growth factor-β1 gene results in multifocal inflammatory disease. Nature 359:693–699

Werner F, Jain MK, Feinberg MW, Sibinga NE, Pellacani A, Wiesel P, Chin MT, Topper JN, Perrella MA, Lee ME (2000) Transforming growth factor-β1 inhibition of macrophage activation is mediated via Smad3. J Biol Chem 275:36653–36658

Yamada Y, Hosoi T, Makimoto F, Tanaka H, Seino Y, Ikeda K (1999) Transforming growth factor beta-1 gene polymorphism and bone mineral density in Japanese adolescents. Am J Med 106:477–479