Structure and organization of the human *TRKA* gene encoding a high affinity receptor for nerve growth factor

Yasuhiro Indo*, Sek Mardy, Motoko Tsuruta, Mohammad Azharul Karim**,
and Ichiro Matsuda

Department of Pediatrics, Kumamoto University School of Medicine,
1-1-1 Honjo, Kumamoto 860, Japan

* To whom correspondence should be addressed.

** Present address: Department of Medical Genetics and Pediatrics, University of Wisconsin-Madison Medical School, Madison, Wisconsin, U.S.A.

(Received January 28, 1997; Revised version accepted April 8, 1997)

Summary

Nerve growth factor (NGF) induces neurite outgrowth and promotes survival of embryonic sensory and sympathetic neurons. TRKA, a receptor tyrosine kinase cloned from a human colon cancer was later found to be expressed in the nervous system and phosphorylated in response to NGF. Somatic rearrangement(s) of the *TRKA* gene (also designated *NTRK1*) are responsible for formation of some oncogenes. Genetic defects in TRKA are responsible for a human disorder, congenital insensitivity to pain with anhidrosis (CIPA). We report here isolation and characterization of the *TRKA* gene which spans at least 23 kb and is split into 17 exons. Exon sizes range from 18 to 394 bp and intron sizes range from 170 bp to at least 3.3 kb. Sizes and boundaries of the exons were determined, and all the splice donor and acceptor sites conformed to the GT/AG rule. Approximately 1.2 kb of the 5'-flanking regions was sequenced, and putative regulatory elements were identified. These results will be useful for studies on the developmental and biological regulation of the *TRKA* gene and for further characterization of mutations in CIPA patients as well as elucidation of mechanisms responsible for rearrangement(s) observed in human tumors.

Key Words

TRKA, nerve growth factor, nerve growth factor receptor, receptor tyrosine kinase, congenital insensitivity to pain with anhidrosis

Cell survival, growth and differentiation in nervous systems are mediated by numerous growth factors, including neurotrophic factors (neurotrophins). Nerve growth factor (NGF), the first neurotrophic factor to be discovered, supports the survival of sympathetic ganglion neurons and subpopulations of mainly nociceptive sensory neurons in dorsal root ganglia derived from the neural crest as well as ascending cholinergic neurons of the basal forebrain (Levi-Montalcini, 1987; Thoenen and Barde, 1980). The *TRKA* (also named *NTRK1*) was isolated from a colon carcinoma as a potential new member of the tyrosine kinase gene family (Martin-Zanca *et al*., 1986) and was later found to be expressed in the nervous system (Martin-Zanca *et al*., 1990). TRKA is a receptor tyrosine kinase and is phosphorylated in response to NGF (Kaplan *et al*., 1991; Klein *et al*., 1991).

Congenital insensitivity to pain with anhidrosis (CIPA; McKusick: 256800, also known as congenital sensory neuropathy with anhidrosis, hereditary sensory and autonomic neuropathy type IV) is an autosomal-recessive genetic disease characterized by recurrent episodes of unexplained fever, anhidrosis (absence of sweating) and absence of reaction to noxious stimuli, self-mutilating behavior, and mental retardation (Swanson, 1963; Dyck, 1984; McKusick, 1994). Recently, we have reported that the gene responsible for CIPA is *TRKA*, suggesting that the NGF-TRKA system plays a crucial role in development and function of the nociceptive reception as well as establishment of thermoregulation via sweating systems in humans (Indo *et al*., 1996).

The *TRKA* gene is located on the q arm of chromosome 1 (Miozzo *et al*., 1990; Morris *et al*., 1991). The *TRKA* -derived oncogenes are also detected in human breast tumor cells (Kozma *et al*., 1988) or in papillary thyroid carcinoma (Butti *et al*., 1995; Greco *et al*., 1995). These oncogenes are activated by somatic rearrangements juxtaposing their tyrosine kinase domain to the 5’-end sequences derived from unrelated loci and producing chimeric oncogenes whose products display a constitutive and ectopic tyrosine kinase activity. Breakpoints producing some oncogenes often involve a specific region of the *TRKA* gene and part of its sequence has been described (Greco *et al*., 1993). However, the exact location of this region in the whole *TRKA* gene is unknown, as structure of this gene has not been documented.

We have now defined the structural and genomic organization of *TRKA*. This knowledge will be useful for studies in developmental and biological regulation of the *TRKA* gene and further characterization of mutation(s) in CIPA patients and as well as elucidation of mechanisms responsible for rearrangement(s) of *TRKA* in human tumors.

A phage library constructed from human leukocytes (Clontech, Palo Alto, CA) was screened to obtain DNA fragments from the *TRKA* gene. We used a human *TRKA* cDNA (pLM6) (Martin-Zanca *et al.*, 1989) as a radioactively labeled probe and isolated two clones (T6 and T11), covering the entire genomic region of *TRKA*, as shown in Fig. 1. We previously characterized the *TRKA* gene encoding the intracellular domain (Indo *et al.*, 1996). In the present study, all the exon/intron splice junctions were determined by comparing the human *TRKA* genomic sequences with the human *TRKA* cDNA sequences. Size of introns was estimated by the sequence of restriction fragments or polymerase chain reaction. The human *TRKA* gene divided into 17 exons ranged in size from 18 bases (exon 9) to 394 bases (exon 17), and 16 introns ranged in size from 170 bases (intron 9) to at least 3.3 kb (intron 1). The entire human *TRKA* gene was estimated to span at least 23 kb. The sequences of exon-intron boundaries are presented in Table 1. All of the splice donor and acceptor sites conformed to the GT/AG rule for nucleotides immediately flanking the exon border (Shapiro and Senapathy, 1987).

A single transmembrane domain divides the TRKA protein into an extracellular and an intracellular domain (Snider, 1994; Barbacid, 1995). The extracellular domain is important for specific NGF binding and includes a signal peptide, three tandem leucine-rich motifs flanked by two cysteine clusters, and two immunoglobulin-like domains (or motifs). The intracellular domain includes a juxtamembrane region, a tyrosine kinase domain, and a very short carboxy-terminal tail (Barbacid, 1995). We found a general correlation between the genomic organization of the *TRKA* gene and the functional organization of TRKA protein. Exon 1 contains the signal peptide and the first cysteine cluster. Three leucine-rich motifs are encoded by exons 2, 3 and 4, respectively. This suggests that simple duplication can account for the variable numbers of the motif. Exon 5 contains the second cysteine cluster. The first immunoglobulin-like motif is encoded by exons 6 and 7, while the second immunoglobulin-

like motif is encoded by the single exon 8. Thus, the splice sites of the TRKA gene encoding the extracellular domain separate the functional domains so that each domain is encoded by separate exons. Exon 9 is a small (18-bp) one incorporated into mRNA by alternative splicing and six amino acid residues encoded by this exon are present in the extracellular domain of the neuronal-specific TRKA receptor (Barker *et al*., 1993). The transmembrane domain is encoded by exons 10 and 11 and the intracellular domain of TRKA is encoded by exons 11-17. The juxtamembrane domain is encoded by exons 11 and 12. The domain contains an IXNPXpY motif where p indicates phosphorylation at Tyr-490 residue of the activated TRKA (Dikic *et al*., 1995). This motif is encoded by exon 12 and is recognized by an Shc adaptor protein required for activation of the Ras-MAPK pathway (Obermeier *et al*., 1994; Stephens *et al*., 1994). The tyrosine kinase domain which is phosphorylated in response to NGF and is critical for the intercellular signaling is encoded by exons 13-17. A consensus sequence motif YXXM which interacts with phosphatidylinositol-3' kinase, is located at the end of the kinase catalytic domain (Tyr-751 residue in TRKA) (Obermeier *et al*., 1993a; Soltoff *et al*., 1992) and encoded by exon 17. The short carboxy-terminal tail of 15 amino acids is also encoded by exon 17 and includes a conserved Tyr residue (Tyr-785 in TRKA) which is responsible for binding of phospholipase Cγ (Obermeier *et al*., 1993b; Loeb *et al*., 1994). Thus, the functional domains or motifs are generally encoded by different exons, except for exons 10 and 11 which contain a small portion of the transmembrane region and a portion of the juxtamembrane region, respectively.

In addition, there were discrepancies between the nucleotide sequences in exons 7 and 8 of the genomic clone (T6) and those noted in the cDNA (Martin-Zanca *et al*., 1989). A single base substitution of T-871 to G in exon 7 and substitutions of dinucleotide CG (983 and 984) to GC in exon 8 changed codons as follows: Leu-263 to Val and Ser-300 to Cys, respectively. We also found these two substitutions in cDNAs from two normal controls and four CIPA subjects (data not shown). These amino acid changes were located in the immunoglobulin-like motifs 1 and 2, respectively. In rat *TRKA* cDNA, two amino acids corresponding to these residues are also Val and Cys, respectively (Meakin *et al*., 1992). It is noteworthy because this immunoglobulin-like motif-2 is the structural element that determines the interaction of

neurotrophins with their receptors (Urfer *et al*., 1995). If our data are accurate, two Cys residues should be conserved in the second immunoglobulin-like motif of all human TRK family members as in the first immunoglobulin-like motif (Nakagawara *et al*., 1995).

The nucleotide sequence of 1226 bp upstream from the base number 1 of the *TRKA* cDNA (Martin-Zanca, *et al*., 1989) is shown in Fig. 2a. A consensus CAAT or TATA elements were not present upstream of the putative region for transcription initiation. To determine regulatory sites in the TRKA gene, we used a computer program "TFSEARCH" based on a database, "TRANSFAC", which compiled eukaryotic *cis*-acting regulatory DNA elements and *trans*-acting factors (Wingender, 1994). Sequences from vertebrates were selected on the threshold of 95.0. Sequences similar to the binding site for several transcription factors located between -420 and -990. Sequences homologous to the binding site for the following proteins were seen in this region: c-Rel /NF-κB (Kunsch *et al*., 1992; Baeuerle and Henkel, 1994; Schmidt *et al*., 1996) at position -898, AP-2 (Mitchell *et al*., 1991; Faisst and Meyer, 1992) at position -852, CdxA (Margalit *et al*., 1993; Frumkin *et al*., 1993) at position -809, CCAAT displacement protein (CDP-CR) (Neufeld *et al*., 1992; Harada *et al*., 1995) at position -767, and heat shock transcription factor (HSF) at position -429 (both directions) (Kroeger and Morimoto, 1994). Determination of the precise putative promoter sites in transcriptional regulation of the TRKA gene is the subject of ongoing study. The nucleotide sequence of the 3' exon region of the gene is shown in Fig. 2b. The site of the polyadenylation signal was inferred from the cDNA. A noncanonical polyadenylation signal was present 21 bases upstream of the polyadenine tail.

Oncogenic rearrangements often involve the same region of the *TRKA* gene, resulting in the same junction. Our study indicates that the region frequently involved in the rearrangements is located in exons 8 through 12 of the *TRKA*. The structural and genomic organization of the whole human *TRKA* gene will provide a basis for elucidation of mechanisms responsible for such rearrangement(s). In addition, we found a microsatellite region (GT or CA repeat) located in intron 12. The nucleotide sequence flanking this locus was reported (Greco *et al*., 1993) and discussed (Butti *et al*., 1995). Position of this locus (AFMa127wh9) and data on the heterozygosity were described (Dib *et al*., 1996).

Clinically, CIPA is a serious illness that might be fatal in the first years of life if the hyperpyrexia is not properly overcome. In older children, osteomyelitis and bone and/or joint deformities demand surgical procedures sometimes involving extensive amputations (Dyck, 1984; McKusick, 1994). To data, three different mutations in four families have been identified. A deletion-, a splice- and a missense-mutations all in the region encoding the tyrosine kinase domain were detected in these families. The present study revealed that a single base deletion and missense mutations are located in exon 14 of this gene. Splice mutation is located in the 5'-splice donor site of intron 15. This study will facilitate analyses of CIPA mutations in other regions of *TRKA.* Identification of such mutations will further genetic diagnoses of this painless but serious disease.

Acknowledgements

We thank L.F. Parada for providing the pLM6 plasmid and M. Ohara for assistance with the manuscript. This work was funded in part by a Grant-in-Aid for Scientific Research (C) from the Ministry of Education, Science, Sports and Culture of Japan and the Research Grant (8A-7) for Nervous and Mental Disorders from the Ministry of Health and Welfare of Japan.

References

Baeuerle PA, Henkel T (1994): Function and activation of NF-κB in the immune system. Annu Rev Immunol **12**: 141-179

Barbacid M (1995): Structural and functional properties of the TRK family of neurotrophin receptors. Ann N Y Acad Sci **766**: 442-458

Barker PA, Lomen HC, Gensch EM, Meakin SO, Glass DJ, Shooter EM (1993): Tissue-specific alternative splicing generates two isoforms of the *trkA* receptor. J Biol Chem **268**: 15150-15157

Butti MG, Bongarzone I, Ferraresi G, Mondellini P, Borrello MG, Pierotti MA (1995): A sequence analysis of the genomic regions involved in the rearrangements between TPM3 and NTRK1 genes producing TRK oncogenes in papillary thyroid carcinomas. Genomics **28**: 15-24

Dib C, Faure S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J (1996): A comprehensive genetic map of the human genome based on 5,264 microsatellites. Nature **380**: 152-154

Dikic I, Batzer AG, Blaikie P, Obermeier A, Ullrich A, Schlessinger J, Margolis B (1995): Shc binding to nerve growth factor receptor is mediated by the phosphotyrosine interaction domain. J Biol Chem **270**: 15125-15129

Dyck PJ (1984): Neuronal atrophy and degeneration predominantly affecting sensory and autonomic neurons. In: Dyck PJ, Thomas PK, Lambert EH, Bunge R (eds). Peripheral neuropathy. W. B. Saunders Company, Philadelphia. pp 1557-1599

Faisst S, Meyer S (1992): Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res **20**: 23-26

Frumkin A, Haffner R, Shapira E, Tarcic N, Gruenbaum Y, Fainsod A (1993): The chicken *CdxA* homeobox gene and axial positioning during gastrulation. Development **118**: 553-562

Greco A, Mariani C, Miranda C, Pagliardini S, Pierotti MA (1993): Characterization of the NTRK1 genomic region involved in chromosomal rearrangements generating TRK oncogenes. Genomics **18**: 397-400

Greco A, Mariani C, Miranda C, Lupas A, Pagliardini S, Pomati M, Pierotti MA (1995): The DNA rearrangement that generates the *TRK-T3* oncogene involves a novel gene on chromosome 3 whose product has a potential coiled-coil domain. Mol Cell Biol **15**: 6118-6127

Harada R, Berube G, Tamplin OJ, Denis LC, Nepveu A (1995): DNA-binding specificity of the cut repeats from the human cut-like protein. Mol Cell Biol **15**: 129-140

Indo Y, Tsuruta M, Hayashida Y, Karim MA, Ohta K, Kawano T, Mitsubuchi H, Tonoki H, Awaya Y, Matsuda I (1996): Mutations in the *TRKA*/NGF receptor gene in patients with congenital insensitivity to pain with anhidrosis. Nature Genet **13**: 485-488

Kaplan DR, Hempstead BL, Martin ZD, Chao MV, Parada LF (1991): The trk proto-oncogene product: a signal transducing receptor for nerve growth factor. Science **252**: 554-558

Klein R, Jing SQ, Nanduri V, O'Rourke E, Barbacid M (1991): The trk proto-oncogene encodes a receptor for nerve growth factor. Cell **65**: 189-197

Kozma SC, Redmond SM, Fu XC, Saurer SM, Groner B, Hynes NE (1988): Activation of the receptor kinase domain of the *trk* oncogene by recombination with two different cellular sequences. EMBO J **7**: 147-154

Kroeger PE, Morimoto RI (1994): Selection of new HSF1 and HSF2 DNA-binding sites reveals difference in trimer cooperativity. Mol Cell Biol **14**: 7592-7603

Kunsch C, Ruben SM, Rosen CA (1992): Selection of optimal κB/Rel DNA-binding motifs: interaction of both subunits of NF-κB with DNA is required for transcriptional activation. Mol Cell Biol **12**: 4412-4421

Levi-Montalcini R (1987): The nerve growth factor: thirty-five years later. EMBO J **6**: 1145-1154

Loeb DM, Stephens RM, Copeland T, Kaplan DR, Greene LA (1994): A Trk nerve growth factor (NGF) receptor point mutation affecting interaction with phospholipase C-γ 1 abolishes NGF-promoted peripherin induction but not neurite outgrowth. J Biol Chem **269**: 8901-8910

Margalit Y, Yarus S, Shapira E, Gruenbaum Y, Fainsod A (1993): Isolation and characterization of target sequences of the chicken CdxA homeobox gene. Nucleic Acids Res **21**: 4915-4922

Martin-Zanca D, Hughes SH, Barbacid M (1986): A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. Nature **319**: 743-748

Martin-Zanca D, Oskam R, Mitra G, Copeland T, Barbacid M (1989): Molecular and biochemical characterization of the human *trk* proto-oncogene. Mol Cell Biol **9**: 24-33

Martin-Zanca D, Barbacid M, Parada LF (1990): Expression of the *trk* proto-oncogene is restricted to the sensory cranial and spinal ganglia of neural crest origin in mouse development. Genes Dev **4**: 683-694

McKusick VA (1994): Mendelian inheritance in man. The Johns Hopkins University Press, Baltimore

Meakin SO, Suter U, Drinkwater CC, Welcher AA, Shooter EM (1992): The rat trk protooncogene product exhibits properties characteristic of the slow nerve growth factor receptor. Proc Natl Acad Sci USA **89**: 2374-2378

Miozzo M, Pierotti MA, Sozzi G, Radice P, Bongarzone I, Spurr NK, Della PG (1990): Human TRK proto-oncogene maps to chromosome 1q32-q41. Oncogene **5**: 1411-1414

Mitchell PJ, Timmons PM, Hebert JM, Rigby PW, Tjian R (1991): Transcription factor AP-2 is expressed in neural crest cell lineages during mouse embryogenesis. Genes Dev **5**: 105-119

Morris CM, Hao QL, Heisterkamp N, Fitzgerald PH, Groffen J (1991): Localization of the TRK proto-oncogene to human chromosome bands 1q23-1q24. Oncogene **6**: 1093-1095

Nakagawara A, Liu XG, Ikegaki N, White PS, Yamashiro DJ, Nycum LM, Biegel JA, Brodeur GM (1995): Cloning and chromosomal localization of the human TRK-B tyrosine kinase receptor gene (NTRK2). Genomics **25**: 538-546

Neufeld EJ, Skalnik DG, Lievens PM, Orkin SH (1992): Human CCAAT displacement protein is homologous to the *Drosophila* homeoprotein, cut. Nature Genetics **1**: 50-55

Obermeier A, Lammers R, Wiesmuller KH, Jung G, Schlessinger J, Ullrich A (1993a): Identification of Trk binding sites for SHC and phosphatidylinositol 3'-kinase and formation of a multimeric signaling complex. J Biol Chem **268**: 22963-22966

Obermeier A, Halfter H, Wiesmuller KH, Jung G, Schlessinger J, Ullrich A (1993b): Tyrosine 785 is a major determinant of Trk-substrate interaction. EMBO J **12**: 933-941

Obermeier A, Bradshaw RA, Seedorf K, Choidas A, Schlessinger J, Ullrich A (1994): Neuronal differentiation signals are controlled by nerve growth factor receptor/Trk binding sites for SHC and PLC γ. EMBO J **13**: 1585-1590

Schmidt UR, Memet S, Lilienbaum A, Feuillard J, Raphael M, Israel A (1996): NF-κB activity in transgenic mice: developmental regulation and tissue specificity. Development **122**: 2117-2128

Shapiro MB, Senapathy P (1987): RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res **15**: 7155-7174

Snider WD (1994): Functions of the neurotrophins during nervous system development: what the knockouts are teaching us. Cell **77**: 627-638

Soltoff SP, Rabin SL, Cantley LC, Kaplan DR (1992): Nerve growth factor promotes the activation of phosphatidylinositol 3-kinase and its association with the trk tyrosine kinase. J Biol Chem **267**: 17472-7

Stephens RM, Loeb DM, Copeland TD, Pawson T, Greene LA, Kaplan DR (1994): Trk receptors use redundant signal transduction pathways involving SHC and PLC-γ 1 to mediate NGF responses. Neuron **12**: 691-705

Swanson AG (1963): Congenital insensitivity to pain with anhidrosis. Arch Neurol **8**: 299-306

Thoenen H, Barde Y-A (1980): Physiology of nerve growth factor. Physiol Rev **60**: 1284-1335

Urfer R, Tsoulfas P, O'Connell L, Shelton DL, Parada LF, Presta LG (1995): An immunoglobulin-like domain determines the specificity of neurotrophin receptors. EMBO J **14**: 2795-2805

Wingender E (1994): Recognition of regulatory regions in genomic sequences. J Biotechnol **35**: 273-280

Figures

Fig. 1. Physical map of the human *TRKA* gene.

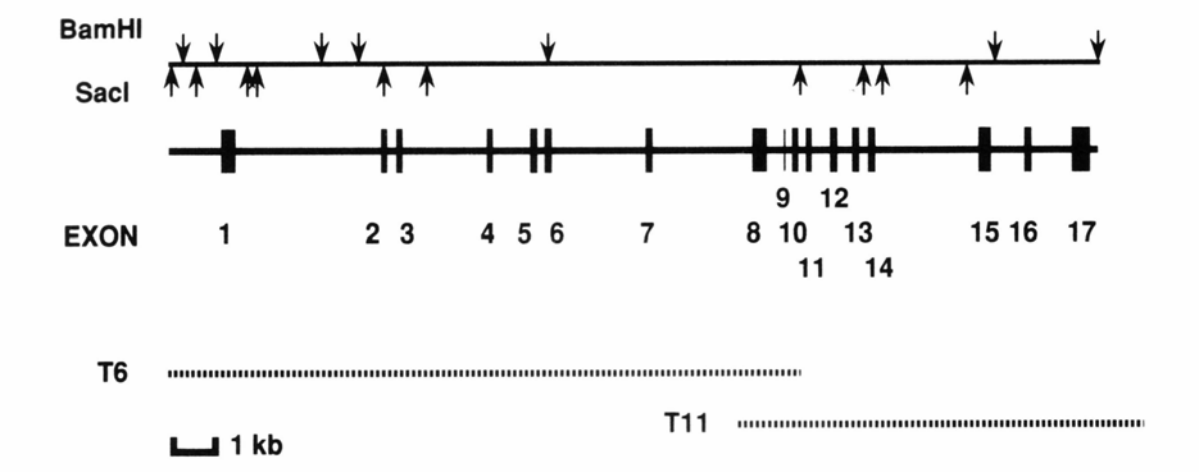

Fig. 1. Physical map of the human *TRKA* gene.

The structure of the gene is shown by a thick line. Exons 1-17 are shown as vertical lines and numbered. Below the gene structure, the genomic DNA fragments from the phage clones are shown by dotted lines. *Bam*HI and *Sac*I sites are shown above.

Fig. 2. Nucleotide sequences of the 5’ and 3’ regions of the human *TRKA* gene.

Fig. 2. Nucleotide sequences of the 5’ and 3’ regions of the human *TRKA* gene.

(a) Nucleotide sequence of the 5’ region of the gene.

The sequence is numbered according to Martin-Zanca et al (Martin-Zanca *et al*., 1989), starting from guanine at base number 1 of the cDNA. Sequences with a horizontal arrow indicate the putative region resembling binding sites for specific transcription factors. The directions of the sequence are indicated by the directions of the arrows. The translation start codon ATG is underlined.

(b) Nucleotide sequences of the 3’ region of the gene.

The termination codon, TAG, is marked (***). The sequence is numbered starting from thymidine of the termination codon. A vertical arrowhead indicates the poly(A) addition site. The noncanonical poly(A) addition signal is boxed.

Table 1

Exon-Intron Organization of the human TRKA gene

Exon	Size (bp)	cDNA position[a]	Intron	Size (kb)	Exon	Intron	Exon
1	296	1-296	1	3.3	GAG CT Glu Leu	gtgagtgtccggcgggcggt ·····	cccatccgctctccccacag C TAC Leu Tyr
2	75	297-371	2	0.3	AAC CT Asn Leu	gtgagggaaacggggactgt ·····	cctcctgcacccctccccag C ACC Leu Thr
3	72	372-443	3	2.1	CGC CT Arg Leu	gtgagtgtggccagtgctgg ·····	ctgtgtctccacgcccgcag G AAT Leu Asn
4	69	444-512	4	1.2	GAA CT Glu Leu	gtgagtgggggcgcttccag ·····	gtgtcccccatgcccccccag G GTC Leu Val
5	146	513-658	5	0.3	TGT G Cys Gly	gtaggtgccgggtgagggag ·····	cgggcgtcctgggtggccag GT GTG Gly GTG
6	143	659-801	6	2.3	GTG ATG Val Met	gtgagaagaccttcgctggc ·····	ccctctctttcctgatctag AAA TCT Lys Ser
7	133	802-934	7	2.1	TCC T Ser Phe	gtgagtctcagtggcagctc ·····	ttgctctttctggcccacag TC CCG Phe Pro
8	327	935-1261	8	0.4	CCT G Pro Val	gtgcgagggccatcctgaac ·····	tctcctccctcctgctgcag TC TCC Val Ser
9[b]	18	-	9	0.2	GTG G Val Asp	gtgagtagcccaaggtggag ·····	cctgccctgtgtccctacag AC ACT Asp Thr
10	56	1262-1317	10	0.3	TTT GGG Phe Gly	gtgagataggaagtagaagc ·····	ctacccctgtcccccaccag GTC TCG Val Ser
11	103	1318-1420	11	0.5	AAC C Asn Arg	gtgagtcggggctgcagaggg ·····	cggctgtgtctcctctctag GC CCG Arg Pro
12	147	1421-1567	12	0.4	GCC T Ala Cys	gtgaggggctatgctgggtc ·····	gaccctgcaagccccctcag GT GTT Cys Val
13	131	1568-1698	13	0.2	GTC AAG Val Lys	gtgagaccctgccccggggg ·····	tccctgccgcttccatccag GCA CTG Ala Leu
14	173	1699-1871	14	2.5	CTC CG Leu Arg	gtaccagcacctggcctcag ·····	tctccttttcttgttcacag A TCC Arg Ser
15	241	1872-2112	15	0.7	TAC CGT Tyr Arg	gtaagggtcctttgtcccca ·····	gcagtgtccgcccgtggcag GTG GGA Val Gly
16	159	2113-2271	16	1.2	ACG GAG Thr Glu	gtcagccccggcccatggtc ·····	tgtctctccggtggccccag GCA ATC Ala Ile
17	394	2272-2665			CTT G		

[a] Numbered according to Martin-Zanca *et al*. (1986).

[b] A small exon incorporated into mRNA of neural tissues by alternative splicing (Barker *et al*., 1993).