Two novel genes in the center of the 11p15 imprinted domain escape genomic imprinting

Maxwell P. Lee¹, Sheri Brandenburg¹, Gregory M. Landes³, Mark Adams⁴, Glenn Miller³ and Andrew P. Feinberg¹,²,*

Departments of ¹Medicine, ²Molecular Biology and Genetics, and Oncology, Johns Hopkins University School of Medicine, 1064 Ross, 720 Rutland Avenue, Baltimore, MD 21205, USA, ³Genzyme Research and Development, One Mountain Road, Framingham, MA 01701, USA and ⁴The Institute for Genome Research, 9712 Medical Center Drive, Rockville, MD 20856, USA

Received November 27, 1998; Revised and Accepted January 17, 1999 DDBJ/EMBL/GenBank accession nos AF125568 and AF125569

We previously reported the isolation of a 2.5 Mb tumor-suppressing subchromosomal transferable fragment (STF) from human chromosome 11p15 and the identification of nine known genes and four novel genes within this STF. We now report the isolation of two novel cDNAs, designated here as TSSC4 and TSSC6 (tumor-suppressing STF cDNA 4 and 6), located within the STF. TSSC4 and TSSC6 encode predicted proteins of 329 and 290 amino acids, respectively, with no close similarity to previously reported proteins. TSSC4 and TSSC6 are both located in the center of a 1 Mb imprinted domain, which contains the imprinted genes TSSC3, TSSC5, p57KIP2, KvLQT1, ASCL2, IGF2 and H19. However, we found that neither TSSC4 nor TSSC6 was significantly imprinted in any of the fetal or extra-embryonic tissues examined. Based on this result, the imprinted gene domain of 11p15 appears to contain at least two imprinted subdomains, between which TSSC4 and TSSC6 substantially escape imprinting, due either to lack of initial silencing or to an early developmental relaxation of imprinting.

INTRODUCTION

Human chromosome 11p15.5 is one of the most heavily studied genomic regions, since it harbors multiple imprinted genes and, thus, may constitute a relatively large imprinted chromosomal domain. In addition, 11p15 harbors several genes, mutations or altered imprinting of which are associated with human disease. These include Beckwith–Wiedemann syndrome (BWS), an autosomal dominant disorder, characterized by prenatal overgrowth and predisposition to cancer; and Wilms’ tumor 2 (WT2), defined by frequent loss of heterozygosity (LOH) in embryonal tumors such as Wilms’ and other embryonal cancers, as well as lung, breast and ovarian cancer. These disorders appear to involve imprinted genes, because BWS is transmitted predominantly through the mother, and balanced germline chromosomal rearrangements in BWS, as well as LOH in embryonal cancers, exclusively involve the maternally inherited allele (reviewed in ref. 1).

We and others have identified seven imprinted genes on 11p15: (i) IGF2, which encodes an important autocrine growth factor in cancer (2–5); (ii) H19, an untranslated RNA whose imprinting regulates IGF2 (6); (iii) ASCL2, a homolog of Drosophila achaete-scute expressed in the trophoblast (7); (iv) KvLQT1, coding for a voltage-gated potassium channel (8); (v) p57KIP2, encoding a cyclin-dependent kinase inhibitor (9); (vi) TSSC5 (also known as IMPT1/BWR1/ORCQL2), which codes for a predicted transmembrane transporter (10–12); and (vii) TSSC3 (also known as JPL) (13,14), a homolog of a mouse apoptosis-inducing gene (15). With the exception of IGF2, all of these genes are expressed from the maternal allele.

We and others have also found direct evidence of a role for five of these genes in disease: (i) IGF2 undergoes loss of imprinting (LOI) in 70% of Wilms’ tumors and about half of all adult cancers, with activation of the normally silent maternal allele (2,5); (ii) H19 shows epigenetic silencing in some tumors with LOI of IGF2 (16); (iii) KvLQT1 is rearranged in most BWS patients with balanced germline chromosomal rearrangements (8); (iv) p57KIP2 is mutated in 5% of BWS patients and shows epigenetic silencing and/or LOI in all Wilms’ tumors (17–21); and (v) TSSC5 undergoes germline mutations in Wilms’ tumor and somatic mutation in lung cancer (11). Because of the large number of imprinted genes on 11p15, spanning ~1 Mb, it is generally believed that 11p15 represents one of two known large (>1 Mb) imprinted domains in the human genome, the other being the Prader–Willi/Angelman syndrome domain of 15q11–13 (22,23). We previously isolated a subchromosomal transferable fragment (STF) which suppresses in vitro growth of the rhabdomyosarcoma cell line RD, confirming the existence of one or more tumor suppressor genes within this region (24). The STF spans ~2.5 Mb, with DIIS12 at its proximal end and DIIS1318 at its distal end (25). Here we report two novel genes, within a cluster of imprinted genes in this STF, that are not imprinted in any of the fetal or extra-embryonic tissues examined.

*To whom correspondence should be addressed. Tel: +1 410 614 3489; Fax: +1 410 614 9819; Email: afeinberg@jhu.edu
Figure 1. Map of TSSC4 and TSSC6 in 11p15.5 and polymorphisms within them. (A) Transcript map. The arrows indicate transcriptional orientation. 161F14 and 318P18 are BACs containing TSSC4 and TSSC6. (B) Intron–exon structure of TSSC4. The polymorphic sites at nucleotides 1290 and 1353 are indicated by a single asterisk because of their proximity. The sequences of primers TSSC4-114, TSSC4-112 and TSSC4-212 are given in Materials and Methods. (C) Partial intron–exon structure of TSSC6. The polymorphic site at nucleotide 89 is indicated by an asterisk. The sequences of primers TSSC6-1A, TSSC6-1B, TSSC6-105 and TSSC6-202 are given in Materials and Methods.

RESULTS

Molecular cloning of TSSC4

We previously reported the mapping and isolation of 13 genes within an STF from human chromosome 11p15.5 (11,25) of which seven were imprinted (2–14). To continue our effort of systematic isolation and characterization of genes from this region, we isolated 12 exons from BAC clones 161F14 and 318P18, which are located between KvLQT1 and IGF2 (Fig. 1A), and are part of a genomic contig that we had isolated earlier (25). This region has been relatively less well characterized than regions located more centromeric or telomeric to this interval (25). As described in Materials and Methods, we identified two expressed sequence tag (EST) contigs within this interval, THC176530 and THC156825. Sequencing of these contigs revealed the presence of a single open reading frame (ORF) encoding a protein of 329 amino acids, within a cDNA of 1463 bp. This gene is designated here as TSSC4 (tumor-suppressing STF cDNA 4; GenBank accession no. AF125568). The first ATG of the ORF is located at nucleotide 182, and the stop codon is at nucleotide 1171 (Fig. 2A). A BLAST search indicated the absence of any protein homologous to TSSC4. Motif analysis using the GCG software package identified one predicted protein kinase A (PKA) site at Ser311. TSSC4 also contained five predicted protein kinase C (PKC) phosphorylation sites, at Ser90, Ser129, Ser146, Ser248 and Ser307 (Fig. 2A). Comparison of TSSC4 and genomic sequence (GenBank accession no. AC000376) indicated that TSSC4 contains two exons, and the entire coding region is located within the second exon (Figs 1B and 2A).

Molecular cloning of TSSC6

In a similar manner, we identified an additional EST contig, THC133845, within this interval. Sequencing of EST clone 235072 within the contig did not reveal a long ORF. However, RT–PCR using primers designed from the 3′ and 3′ end sequences of 235072, followed by direct sequencing of the RT–PCR product, revealed an ORF of 290 amino acids, within a cDNA of 1241 bp (Fig. 2B). This gene was termed TSSC6 (tumor-suppressing STF cDNA 6; GenBank accession no. AF125569). Sequence comparison of TSSC6 and genomic sequence (GenBank accession no. AC002536) indicated that TSSC6 contains nine exons (Fig. 1C). EST clone 235072 lacks exon 4, causing a shift in the reading frame, most likely because of aberrant RNA splicing (26). The first ATG is located at nucleotide 146, and the stop codon TGA is located at nucleotide 1018 (Fig. 2B). A Psi-BLAST search revealed that TSSC6 shows weak similarity to the rat TAPA1 gene (E = 0.019), which belongs to the tetramembrane-spanning protein family. Interestingly, the human TAPA1 gene is located 60 kb centromeric to TSSC6, suggesting that one of the genes may have arisen in part by duplication followed by considerable sequence divergence (Fig. 1A). Motif analysis indicates the presence of one potential PKA site at Ser18, three potential casein kinase sites at Thr53, Thr87 and Thr138, three potential PKC sites at Thr53, Thr208 and Ser248, and one potential tyrosine kinase phosphorylation site at Tyr24 (Fig. 2B). Mapping of TSSC4 and TSSC6 by Southern hybridization to our previously established genomic contig (25) indicated the following gene order: cen–p57KIP2–KvLQT1–TSSC4–TAPA1–TSSC6–ASCL2–TH–IGF2–tel. The transcriptional orientation for TSSC4, TAPA1 and TSSC6 is from telomere to centromere (Fig. 1A).
Figure 2. DNA and predicted protein sequence of (A) TSSC4. The first methionine of the ORF is at nucleotide 182 and the stop codon is at nucleotide 1171. (B) TSSC6. The first methionine of the ORF is at nucleotide 146 and the stop codon is at nucleotide 1018. The arrows indicate the positions of intron–exon boundaries.

Tissue-specific expression of TSSC4 and TSSC6

To assess the pattern of tissue-specific expression of TSSC4, we performed northern blot hybridization using human multiple tissue northern blots. We detected a 1.6 kb transcript of TSSC4 in fetal brain, lung, liver and kidney (Fig. 3 A), suggesting that the isolated cDNA is full length. TSSC4 was also expressed to varying degrees in all adult tissues (Fig. 3 B–D).

We were not able to detect a signal using TSSC6 on the same MTN blots, suggesting that TSSC6 is expressed at low levels. However, we were able to amplify TSSC6 by RT–PCR from all tissues tested, suggesting ubiquitous low level expression. Since both TSSC4 and TSSC6 are located within the 11p15 imprinted gene domain, we next examined allele-specific expression of TSSC4 and TSSC6. The individual exons of each gene were sequenced from 10 separate fetal samples, using four primer sets to examine the two exons of TSSC4 (one set for exon 1 and three sets for exon 2), and nine primer sets to examine the nine exons of TSSC6. Two transcribed polymorphisms were identified within exon 2 of TSSC4, at nucleotides 1290 and 1353. A transition at nucleotide 1290 from G to A removes a restriction endonuclease site for BanI, while a transition at nucleotide 1353

TSSC4 and TSSC6 were not imprinted in any tissues examined

Since both TSSC4 and TSSC6 are located within the 11p15 imprinted gene domain, we next examined allele-specific expression of TSSC4 and TSSC6. The individual exons of each gene were sequenced from 10 separate fetal samples, using four primer sets to examine the two exons of TSSC4 (one set for exon 1 and three sets for exon 2), and nine primer sets to examine the nine exons of TSSC6. Two transcribed polymorphisms were identified within exon 2 of TSSC4, at nucleotides 1290 and 1353. A transition at nucleotide 1290 from G to A removes a restriction endonuclease site for BanI, while a transition at nucleotide 1353
from C to T adds a HinIII site. We thus used HinIII and BanI to type 35 fetal DNA samples. Four of these were heterozygous (data not shown), which was also confirmed directly by DNA sequencing (Fig. 4).

In order to assess the imprinting status of TSSC4, we carried out RT–PCR using primer pair TSSC4-112/TSSC4-212, generating a 235 bp cDNA product from exon 2 (Fig. 1B), which includes both the G1290A and C1353T polymorphisms. We then directly sequenced the RT–PCR products derived from 15 tissues of four fetuses, to analyze allele-specific gene expression. No RT–PCR products were detected when RT reactions were carried out in the absence of reverse transcriptase, indicating that there was no genomic DNA contamination. These results were also confirmed using a polymorphic HinIII site at nucleotide 1353. Fetuses 1 and 2 were heterozygous (Fig. 6). RT–PCR followed by HinIII digestion revealed biallelic expression of TSSC4 in all eight fetal tissues analyzed. Quantitation using a Phosphorimager indicated that the ratios of a to b alleles, normalizing for fragment size, using genomic DNA as a standard, were 1.22, 0.90, 1.05, 1.04, 1.07, 0.90, 1.07 and 1.37 for adrenal gland, kidney, intestine, lung, liver, heart, lung and limb, respectively (Fig. 6). The absence of a PCR product in the –RT lanes indicated the absence of DNA contamination. These results are consistent with those obtained by DNA sequencing (Fig. 4).

We next examined the imprinting status of TSSC6. Using primer pair TSSC6-1A/TSSC6-1B to amplify exon 1 from genomic DNA (Fig. 1C), we identified a polymorphism changing G at nucleotide 89 to A. We typed 27 fetuses and found seven heterozygotes. We then used primer TSSC6-105 located in exon 4 as the reverse primer for RT–PCR analysis (Fig. 1C), analyzing 35 tissues of four separate fetuses. All 35 tissues displayed biallelic expression. For example, DNA from fetus 5 showed equal representation of both G and A alleles at nucleotide 89 (Fig. 5A, and data not shown). Fetus 8 also showed nearly equal biallelic expression in fetuses 6 and 7, including the kidney, lung, pancreas, adrenal, spleen, eye, thymus, skin, gut, heart, limb, liver and brain (Fig. 5B and C, and data not shown). Fetus 8 also showed nearly equal biallelic expression in most tissues (Fig. 5D, and data not shown), including the kidney, lung, pancreas, adrenal, spleen, eye, thymus, skin, gut, heart, limb, liver and brain (Fig. 5B and C, and data not shown). Fetus 8 also showed nearly equal biallelic expression in most tissues (Fig. 5D, and data not shown), although there was an ∼1.3-fold ratio of expression paternal allele over maternal allele in the lung and brain. This also is in marked contrast to the ratio of expression of other imprinted genes in this region. Finally, examination of two placentas and one amnion showed approximately equal expression of the two alleles, equal to genomic DNA, indicating that TSSC4 is not significantly imprinted in extra-embryonic tissue.

**DISCUSSION**

We have molecularly cloned two novel genes, TSSC4 and TSSC6, within a cluster of imprinted genes on 11p15. We have established the gene order as follows: cen–hNAP2–TSSC3–TSSC5–p57KIP2–Kuffer–LQT1–TSSC4–TAPAI–TSSC6–ASCL2–TH–IGF2–H19–L23–tel. TSSC4 and TSSC6 both lie within a tumor-
Figure 4. Biallelic expression of TSSC4 in tissues of four separate fetuses (A–D). DNA indicates genomic DNA. The sources of RNA for RT–PCR are indicated. The C1353T polymorphism was used for analysis of allele-specific expression in (A) and (B), and the G1290A polymorphism was used in (C) and (D).

suppressing STF from 11p15 (24) and thus are potential tumor suppressor genes.

These results have important implications for the understanding of genomic imprinting. TSSC3, TSSC5, p57KIP2, KvLQT1, ASCL2, IGF2 and H19 previously were shown to be imprinted (2–14). However, we demonstrate here that despite their location in the middle of these genes, both TSSC4 and TSSC6 were not significantly imprinted in any of the fetal or extra-embryonic tissues examined. The ratio of allele-specific expression measured quantitatively was 0.90- to 1.37-fold. While there may be a slight allelic preference in some tissues, TSSC4 and TSSC6 are not substantially imprinted when compared with other genes in this region (4- to >20-fold allelic ratio). This relaxation could arise either by escape from imprinting or by a developmental relaxation of imprinting specific to this region and occurring before day 53 of gestation, when the earliest fetuses could be examined. Consistent with the latter idea, the mouse homolog of TAPA1, which lies between TSSC4 and TSSC6 in humans, has been shown to be imprinted until e8.5, after which imprinting is relaxed (27). Thus, TSSC4 and TSSC6 might also show early developmental relaxation of imprinting, which might have already occurred in the human fetuses. However, mouse and human imprinting are not necessarily the same, as mouse KvLQT1 also loses imprinting after e9.5 (27,28), but the human gene
remains imprinted throughout prenatal development except in the heart (8).

Regardless of whether an imprint fails to include these genes initially or is erased early in development, these data imply the existence of separately regulated subdomains within the 11p15 imprinted domain. This may explain why germline mutations of mouse H19 disrupt Igf2 but have no effect on more distant imprinted genes in cis, such as Kvlqt1 or Mash2 (the mouse homolog of ASCL2) (27).

The results described here predict that the two imprinted subdomains, separated by TSSC4–TSSC6, may define subtypes of human disease. Preliminary data from our laboratory suggest that this is so, as we have found LOI of a novel paternally expressed transcript, termed LIT1, within 11p15 in most patients with BWS (M.P. Lee, M.R. DeBaun, K. Mitsuya, H.L. Galonek, S. Brandenburg, M. Oshimura and A.P. Feinberg, submitted for publication) and LOI of LIT1 in this region is independent of LOI in the more telomeric region including IGF2 (M.P. Lee et al., submitted for publication). Conversely, we and others previously showed frequent LOI of IGF2 in Wilms’ tumor (2,5), and tumors with LOI of IGF2 show normal imprinting of LIT1 in the centromeric subdomain (K. Mitsuya, M. Meguro, M.P. Lee, A. Kashiwagi, M. Kohda, M. Katoh, T.C. Schultz, H. Kugoh, M.A. Yoshida, A.P. Feinberg and M. Oshimura, submitted for publ...
ET9 (0.5 M Tris–HCl pH 9.0, 20 mM EDTA, 10 mM NaCl), at –135

Normal fetal tissues (53–81 days gestation) were obtained from

**Isolation of DNA and RNA from tissues**

The uncut and cut alleles correspond to C1353 (labeled a) and T1353 (labeled b), respectively. The following tissues in fetuses 1 (cDNA lanes 10–17, genomic DNA lane 18) and 2 (cDNA lanes 1–8, genomic DNA lane 9) were examined, in sets of four (–RT followed in the figure by +RT of the same tissues): liver, heart, lung and limb (fetus 1); adrenal, kidney, intestine and limb (fetus 2). Phosphorimager analysis showed an allele ratio (normalized for DNA content) of 0.90 to 1.37.

**MATERIALS AND METHODS**

**cDNA cloning and mapping**

We previously isolated 30 PAC, P1 and BAC clones spanning STF 74-1-6, from which 200 non-redundant trapped exons were ultimately was cloned, as described in Results. Two imprinted subdomains may be more specific for BWS (M.P. Lee et al., submitted for publication).

**Identification of transcribed polymorphisms in TSSC4 and TSSC6**

In order to identify transcribed polymorphisms in both genes, we sequenced both exons of TSSC4, using a total of four sets of primers, on each 10 fetal DNA samples as templates. Similarly, all nine exons of TSSC6 were sequenced using nine sets of primers. As described in Results, two polymorphisms were identified in TSSC4 and one in TSSC6. Primers used for typing the G1290A and C1353T polymorphisms in exon 2 of TSSC4 are as follows: TSSC4-112, 5’-GAGATGGCCAGCTGACCC-3’; and TSSC4-212, 5’-AACCTTTAATTGCTCC-TACAGGGAGCC-3’. Primers used for typing the G89A polymorphism in exon 1 of TSSC6 are as follows: TSSC6-1A, 5’-GCTGTCACTGGCCCTCTGAGCAA-3’; and TSSC6-1B, 5’-TACACTTCTGAGGGGAGC-3’. PCR reactions contained 0.5 µM primers, 0.2 mM dNTP, 50 ng of DNA, 1x PCR buffer (LTI, Gaithersburg, MD), 1.5 mM MgCl2 and 1 U of Taq DNA polymerase (LTI), in 25 µl, and they were performed with a Robocycler (Stratagene, La Jolla, CA) as follows: 40 cycles of 95°C for 45 s, 60°C for 30 s and 72°C for 1 min 30 s; followed by extension at 72°C for 10 min. PCR products were purified using Qiagen II (Qiagen, Valencia, CA) and directly sequenced, or treated with HindIII or BanI and analyzed by electrophoresis on 5% polyacrylamide gels.

**Northern blot hybridization**

Northern blot experiments were performed as described (16) on human multiple tissue northern blots (fetal MTN II, adult MTN, adult MTN III and adult IV; Clontech, Palo Alto, CA). DNA probes were synthesized by the random priming method (30). Autoradiograms were exposed for 24 h.

**Analysis of allele-specific expression in fetal tissues**

RT–PCR and sequencing were used to analyze allele-specific gene expression. RNA was treated with RNase-free DNase (Boehringer Mannheim, Indianapolis, IN) and extracted with phenol–chloroform before the RT reaction. RT was carried out using AMV reverse transcriptase (Boehringer Mannheim). Primers used for RT–PCR are: TSSC4-112 and TSSC4-212 as given in the previous section; TSSC4-114, 5’-AACCTTTAATTGCTCC-TACAGGGAGCC-3’. PCR reactions contained 0.5 µM primers, 0.2 mM dNTP, 50 ng of DNA, 1x PCR buffer (LTI, Gaithersburg, MD), 1.5 mM MgCl2 and 1 U of Taq DNA polymerase (LTI), in 25 µl, and they were performed with a Robocycler (Stratagene, La Jolla, CA) as follows: 40 cycles of 95°C for 45 s, 60°C for 30 s and 72°C for 1 min 30 s; followed by extension at 72°C for 10 min. PCR products were purified using Qiagen II (Qiagen, Valencia, CA) and directly sequenced, or treated with HindIII or BanI and analyzed by electrophoresis on 5% polyacrylamide gels.
REFERENCES


22. Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Aligning DNA sequences. In humans there are relatively few cases of Beckwith–Wiedemann syndrome.
