

## Identical Variant *TSG101* Transcripts in Soft Tissue Sarcomas and Various Non-neoplastic Tissues

Frank Willeke,<sup>1</sup> Ruediger Ridder,<sup>1</sup> Peer Bork,<sup>2</sup> Ruediger Klaes,<sup>1</sup> Gunhild Mechtersheimer,<sup>3</sup> Matthias Schwarzbach,<sup>4</sup> Dagmar Zimmer,<sup>1</sup> Matthias Kloor,<sup>1</sup> Thomas Lehnert,<sup>5</sup> Christian Herfarth,<sup>4</sup> and Magnus von Knebel Doeberitz<sup>1\*</sup>

<sup>1</sup>Division of Molecular Diagnostics and Therapy, Department of Surgery, University of Heidelberg, Heidelberg, Germany

<sup>2</sup>European Molecular Biology Laboratory, Heidelberg, Germany

<sup>3</sup>Institute of Pathology, University of Heidelberg, Heidelberg, Germany

<sup>4</sup>Department of Surgery, University of Heidelberg, Heidelberg, Germany

<sup>5</sup>Division of Surgical Oncology, Department of Surgery, University of Heidelberg, Heidelberg, Germany

Inactivation of the *TSG101* gene was recently shown to induce malignant transformation of NIH/3T3 fibroblasts. Abnormal *TSG101* transcription profiles were observed in various human cancers, and large intragenic deletions of the *TSG101* gene were reported for a series of human breast cancer specimens, pointing to a potential tumor-suppressive activity of *TSG101*. However, subsequent more detailed studies on a large panel of breast carcinoma samples did not confirm the tumor-associated genomic deletions. Here we analyzed the transcription patterns of the *TSG101* gene in soft-tissue sarcomas and non-neoplastic human tissues. Forty-five of 71 soft tissue sarcoma samples (63%) displayed variant transcripts; however, identical aberrant transcripts were also detected in seven of 15 non-neoplastic control tissues. Restriction fragment length polymorphism analysis of the *TSG101* gene excluded major genomic rearrangements in the soft tissue sarcoma samples. Northern blot analysis revealed a very low abundance of variant transcripts as compared with the wild-type *TSG101* transcript. These data point to aberrant splicing of the *TSG101* mRNA in normal and transformed human mesenchymal tissues rather than tumor specific alterations of the *TSG101* gene. In summary, this analyses does not support a pathogenic role for altered *TSG101* expression in human soft tissue sarcomas. *Mol. Carcinog.* 23:195-200, 1998. © 1998 Wiley-Liss, Inc.

Key words: mesenchymal tumors; aberrant *TSG101* transcription; genomic *TSG101*-specific PAC clones

### INTRODUCTION

The *TSG101* gene was identified by a random gene knockout approach in which homozygous inactivation resulted in cellular transformation of NIH/3T3 fibroblasts [1]. Inoculation of these *TSG101*-deficient cells into nude mice led to local tumor formation and distant metastasis. The human homologue of the murine *TSG101* gene was recently cloned and mapped to chromosome 11p15.1-15.2 [2], a genomic locus frequently affected by allelic imbalance in various tumor entities [3-5]. Detailed analyses of the murine and human *TSG101* protein sequences suggested a close relationship to a class of apparently inactive homologues of ubiquitin-conjugating enzymes (E2) that might influence E2 activity in cell-cycle control in a dominant negative way [6,7].

Initially the human *TSG101* gene was reported to be frequently mutated in breast cancer specimens [2]. In a small series of primary human breast cancer samples, genomic long-distance polymerase chain reaction (PCR) experiments suggested large intragenic deletions that were not detected in corresponding normal tissue samples obtained from the same patients. These genomic rearrangements were associated with the generation of truncated *TSG101* transcripts in about 50% of the breast cancer speci-

mens [2]. However, in two subsequent studies, Southern blot analyses of genomic DNA derived from tumor samples did not reveal genomic rearrangements of the *TSG101* gene, although these tumors expressed variant *TSG101* mRNA species [8,9]. Most recently, other tumor types (e.g., ovarian and prostate carcinomas) were also found to harbor variant transcripts [10,11]. Parallel investigations of normal lymphocytes and lung parenchyma revealed the occurrence of variant transcripts as well [10], although the aberrant transcript patterns differ from those found in malignancies. Therefore, variable transcription of the *TSG101* gene seemed not to be restricted to a specific tumor type. Thus, we were interested in whether this gene was altered in mesenchymal malignancies. We characterized the *TSG101* transcription pattern in human soft tissue sarcoma samples and matched

The first two authors contributed equally to this manuscript.

\*Correspondence to: Division of Molecular Diagnostics and Therapy, Department of Surgery, University of Heidelberg, Im Neuenheimer Feld 110, D-69120 Heidelberg, Germany.

Received 25 February 1998; Revised 18 May 1998; Accepted 22 May 1998

Abbreviations: PCR, polymerase chain reaction; RT, reverse transcription; RFLP, restriction fragment length polymorphism; PAC, P1 artificial chromosome.

control tissues by reverse transcription (RT) and subsequent PCR amplification of the cDNAs. Restriction fragment length polymorphism (RFLP) analyses of genomic DNA samples from both various sarcomas and matched non-neoplastic control tissues were performed to investigate whether large genomic rearrangements might be responsible for the generation of truncated *TSG101* transcripts in soft tissue sarcomas. To analyze the actual abundance of variant transcripts in RNA samples, northern blot hybridization experiments were performed.

## MATERIALS AND METHODS

### Tumor and Normal Tissue Samples

Seventy-one tumor samples were investigated from 66 patients surgically treated for soft tissue sarcoma at the Department of Surgery, University of Heidelberg. All tumors were morphologically classified according to the World Health Organization classification [12]. Immediately after surgical removal, the tumor specimens were shock frozen in liquid nitrogen. Twenty-micrometer-thick cryosections with at least 70% tumor cells (as evidenced in control cryostat sections) were used for the extraction of both RNA and DNA. Representative normal tissues (muscle, skin, stomach, and testis) and peripheral venous blood and bone-marrow aspirates were obtained from 15 sarcoma patients. Additionally, the liposarcoma cell line 1955/91 [13] was analyzed.

### Extraction of RNA and *TSG101*-Specific RT-PCR

Total RNA was prepared as described recently [14]. In brief, RNA was extracted from the cell line and from 12–15 consecutive cryosections (20  $\mu\text{m}$  thick) of tumor and normal tissue samples by using a commercially available RNA isolation kit (Glassmax; GIBCO BRL Life Technologies, Karlsruhe, Germany). To eliminate contaminating DNA within RNA preparations, samples were digested with RNase-free DNase I as recommended by the supplier (GIBCO BRL Life Technologies) for 15 min at 25°C.

*TSG101*-specific RT-PCR was performed as described previously [2]. One microgram of total RNA was reverse transcribed with Superscript II (GIBCO BRL Life Technologies) in 20  $\mu\text{L}$  of the reaction mixture by using a *TSG101*-specific primer (5'-ATTTAGCAGTCCCAACATTCAGCACAAA-3') for 60 min at 37°C. Subsequently, 2.5  $\mu\text{L}$  of first-strand cDNA preparation was used for amplification in PCRs containing 25 pmol of both primers P1 (5'-CGGGTGTCTGGAGAGCCAGCTCAGAAA-3') and P2 (5'-CCTCCAGCTGGTATCAGAGAAGTCGT-3'), 200  $\mu\text{M}$  each dNTP, 2.5 U of Taq DNA polymerase (GIBCO BRL Life Technologies), and 1.5 mM  $\text{MgCl}_2$  in a final volume of 50  $\mu\text{L}$  in an Omnigene ThermoCycler (Hybaid, Teddington, UK). After initial template denaturation (95°C for 3 min), amplification was performed for 25 cycles at 95°C

for 30 s, 65°C for 30 s, and 72°C for 1 min and a final extension at 72°C for 5 min. An aliquot (2.5  $\mu\text{L}$ ) of the first PCR products was further amplified by 30 cycles of a nested PCR with primers P3 (5'-AGCCAGCTCAAGAAAATGGTGTCCAAG-3') and P4 (5'-TCACTGAGACCGGCAGTCTTCTTGCTT-3') and identical PCR conditions. All PCRs were repeated at least twice with independent RNA preparations. The PCR products were visualized under ultraviolet light (280 nm wavelength) after electrophoresis in 1% agarose gels stained with ethidium bromide (0.5  $\mu\text{g}/\text{mL}$ ). The specificity of the PCR products was confirmed by Southern blot hybridization with a cDNA probe comprising nt 141–1214 of the *TSG101* cDNA sequence (GenBank entry U82130) and a chemoluminescence-based labeling and detection system (ECL; Amersham Life Science, Amersham, Buckinghamshire, UK).

### Northern Blot Analysis

Ten micrograms of total RNA from sarcoma and normal tissue samples was separated by agarose gel electrophoresis and blotted onto a Hybond N<sup>+</sup> membrane (Amersham Life Science). A *TSG101*-specific cDNA probe comprising nt 106–1482 was radioactively labeled with [ $\alpha$ -<sup>32</sup>P]dATP and a Random Priming DNA labeling kit (MBI Fermentas, Vilnius, Lithuania). Membrane filters were hybridized according to the protocol described by Church and Gilbert [15].

### DNA Sequencing

The PCR products were ligated into vector pCR2.1 (Invitrogen, San Diego, CA) and subjected to nucleotide sequence analysis with a Cy5-AutoRead sequencing kit (Pharmacia Biotech, Freiburg, Germany). The sequencing reaction products were analyzed on denaturing 6.6% polyacrylamide/7 M urea gels with an ALFexpress DNA sequencing device (Pharmacia Biotech).

### Extraction of Genomic DNA and Southern Blot Analysis

Genomic DNA extracted by a standard phenol-chloroform method [16] from leiomyosarcoma, neurogenic sarcoma, and liposarcoma specimens and from non-neoplastic control tissues (muscle, kidney, and testis) was analyzed for RFLPs. Genomic DNA (7.5  $\mu\text{g}$ ) was digested for 12 h with 20 U of the restriction endonucleases HindIII or BglII (MBI Fermentas). After separation of the resulting DNA fragments by 0.9% agarose gel electrophoresis, the DNA digests were blotted onto a Hybond N<sup>+</sup> membrane. The radiolabeled *TSG101*-specific cDNA probe used in the Northern blot experiments was used for Southern blot hybridization of genomic DNA as well. Membrane filters were hybridized for 16 h at 65°C in 0.5 M sodium phosphate, pH 7.2, and 7% sodium dodecyl sulfate and subsequently washed at 65°C in 0.5 $\times$  standard saline citrate and 0.1% sodium dodecyl

sulfate. For the identification of P1 artificial chromosome (PAC) clones containing the *TSG101* gene, high-density filters containing a library of genomic DNA from a human male fibroblast cell line (RPC11-4 human PAC, library no. 704, Resource Center/Primary Database of the German Human Genome Project, Berlin, Germany) ligated into pCYPAC-2 [17] were screened by hybridization as described above. Positive PAC clones LLNLP704P10234Q13 and LLNLP70-4A21138Q13 were obtained from the Resource Center of the German Human Genome Project [18].

## RESULTS

Aberrant *TSG101* transcripts were detected by RT-PCR in 45 of 71 tumor samples (63%) and in liposarcoma cell line 1955/91, all of which showed wild-type *TSG101* cDNA at the same time (Figure 1). In each of the analyzed histological soft tissue sarcoma types, truncated *TSG101* transcripts were discovered (Table 1). Ten of 45 tumor samples with variant PCR products showed more than a single truncated transcript (see Figure 1, second, fourth, and twelfth lanes). However, transcripts different from wild-type *TSG101* were not restricted to malignant lesions. The examination of 15 matched non-neoplastic control tissues revealed variant *TSG101* transcripts in seven samples of peripheral blood monocytes, muscle, testis, and stomach (Figure 1 and Table 1). In general, the abundance of the aberrant transcripts differed in the individual samples. The reproducibility of the aberrant transcription patterns was confirmed by repeated RT-PCRs using independent RNA extracts from all samples. The specificity of all variant PCR fragments was verified by Southern blot hybridization by using a cDNA fragment containing most of the *TSG101* open reading frame but excluding the primer sequences used for RT-PCR amplification (Figure 1).

DNA sequence analysis of all variant types of PCR fragments obtained in this study revealed seven aberrant *TSG101* transcripts (Figure 2 and Table 2), of which type A (247 bp) appeared to be the most abun-

Table 1. Frequency of Variant *TSG101* Transcripts in Soft Tissue Sarcomas of Different Histologies and in Non-Neoplastic Control Tissues

Sample type	No. with variant/ no. examined (%)
Liposarcoma	26/34 (76%)
Synovial sarcoma	5/10 (50%)
Malignant schwannoma	4/6 (66%)
Leiomyosarcoma	2/6 (33%)
Stromal sarcoma	3/5 (60%)
Various sarcoma types	5/10 (50%)
Total sarcoma samples	45/71 (63%)
Normal tissues	7/15 (46%)
Cell line 1955/91 (liposarcoma)	1/1 (100%)

dant in individual samples. Besides aberrant transcripts of types A, B, and D [8], four additional transcripts that have not been reported so far (referred to as types G, H, I, and K) were identified (Figure 2). The variant fragments obtained in RT-PCR analysis of normal tissues in this study included types A, B, D, and G, which were also observed in tumor samples, and also transcript type I, amplified only from a normal muscle biopsy specimen. In contrast to the most frequently detected transcripts, types A and B (Table 2), transcripts D, G, H, I, and K did not include sequences at the junctions of the deleted part of the *TSG101* cDNA that might be genuine or cryptic splice donor or acceptor sites.

To analyze whether genomic deletions contribute to the generation of truncated *TSG101* transcripts, DNA preparations from six soft tissue sarcomas and matched non-neoplastic control tissues (kidney and muscle) were analyzed by Southern blot hybridization using various restriction endonucleases. The normal tissue samples and all of the analyzed soft tissue sarcoma samples had identical hybridization patterns (Figure 3), although truncated *TSG101* transcripts were revealed by RT-PCR in the tumor samples (compare to first six lanes in Figure 1). The hybridization patterns of the genomic DNA samples extracted from normal and neoplastic tissue samples were confirmed by control hybridization of two PAC clones isolated by screening a human PAC library with the *TSG101*-specific DNA probe. Comparable patterns were obtained from purified DNA from PAC clone LLNLP704P10234Q13 digested either with HindIII or BglII (Figure 3), although the HindIII-digested PAC DNA lacked a 1.2-kb fragment and had a larger DNA fragment (7.2 kb) because of the fusion of human DNA fragments to PAC vector sequences at adjacent HindIII recognition sites.

Although we used a nested RT-PCR protocol for the amplification of the *TSG101* transcripts, the relative abundance of several of the obtained PCR fragments was rather low compared with the abundance of wild-type *TSG101* transcript in the same samples.

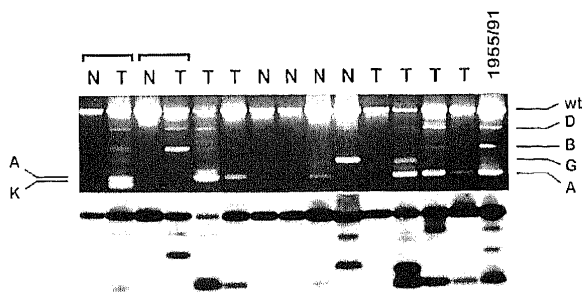


Figure 1. RT-PCR based analysis of variant *TSG101* transcripts in soft tissue sarcomas and matched non-neoplastic control tissues. Truncated transcripts were obtained both in normal tissue samples (N) and in soft tissue sarcoma specimens (T) as well as in the liposarcoma cell line 1955/91. The transcript patterns visible in ethidium bromide-stained agarose gels (upper panel) were confirmed to be specific for *TSG101* by Southern blot analysis (lower panel) by using a chemiluminescence-based protocol for detection. wt, wild-type *TSG101*; A, B, D, G, and K, variant *TSG101* transcripts described in Figure 2.

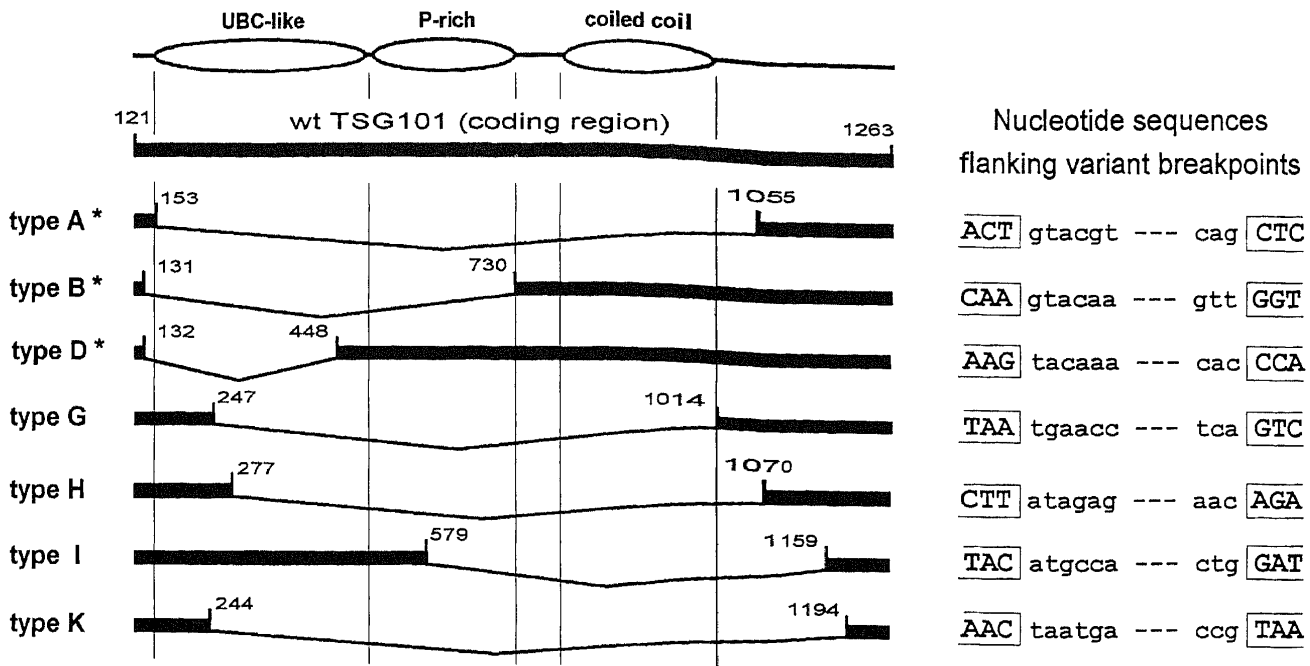


Figure 2. Schematic presentation of wild-type (wt) *TSG101* coding region and variant transcripts (types A, B, D, G, H, I, and K) observed in soft tissue sarcomas and matched normal tissue samples. The proposed organization of the *TSG101* functional domains [7] is indicated. Transcripts previously described by Lee and Feinberg [8] are marked with asterisks. The numbers indi-

cate the nucleotides directly adjacent to the deleted *TSG101* sequence (enumeration according to GenBank entry U82130). UBC, catalytic domain of ubiquitin-conjugating enzymes; P-rich, proline-rich region; coiled coil, coiled-coil domain. Nucleotide sequences flanking the variant breakpoints are given.

In general, variant transcripts were relatively weak RT-PCR DNA fragments, but some variants, mainly variant transcript type A, were present as strong RT-PCR products in most cases. However, this observation might be due to the preferential amplification of smaller DNA fragments during the nested RT-PCR protocol. Therefore, we evaluated the real abundance of variant *TSG101* transcripts by northern blot analysis. Total RNA preparations from various sarcoma and normal tissue samples and from liposarcoma cell line 1955/91 were hybridized with a *TSG101*-specific cDNA probe. Although all of the analyzed samples showed variant transcripts by RT-PCR analysis, by northern blot analysis we detected only one major band, suggesting that wild-type *TSG101* but no variants were expressed (Figure 4).

#### DISCUSSION

We demonstrated that the generation of aberrant transcripts of the human tumor susceptibility gene

*TSG101* was not restricted to epithelial cancers but was also frequent in all histological subtypes of human soft tissue sarcomas. However, identical aberrant transcripts were also shown to occur with comparable frequencies in normal control tissues. These observations suggest that alternative or aberrant splicing events may cause the variant patterns of *TSG101* transcripts found in malignant and nonmalignant tissues of epithelial and mesenchymal origin. A recent report on altered *TSG101* expression profiles in normal blood lymphocytes and lung parenchyma supports this notion, although in that study the observed patterns of *TSG101* transcripts differed between benign and malignant samples [10].

Detailed analyses of the DNA sequences flanking the variant breakpoints revealed sequences resembling genuine or cryptic splice donor and acceptor sites in the two variant transcripts appearing most

Table 2. Variant *TSG101* Transcripts in Normal Tissues and Tumor Samples

Variant <i>TSG101</i> transcript	No. with variant/no. examined (%)	
	Normal tissues	Tumor samples
Type A	1/15 (6.7%)	18/71 (25.4%)
Type B	3/15 (20%)	26/71 (36.6%)
Type D	1/15 (6.7%)	4/71 (5.6%)
Type G	1/15 (6.7%)	5/71 (7%)
Type H	0/15 -	1/71 (1.4%)
Type I	1/15 (6.7%)	0/71 -
Type K	0/15 -	1/71 (1.4%)

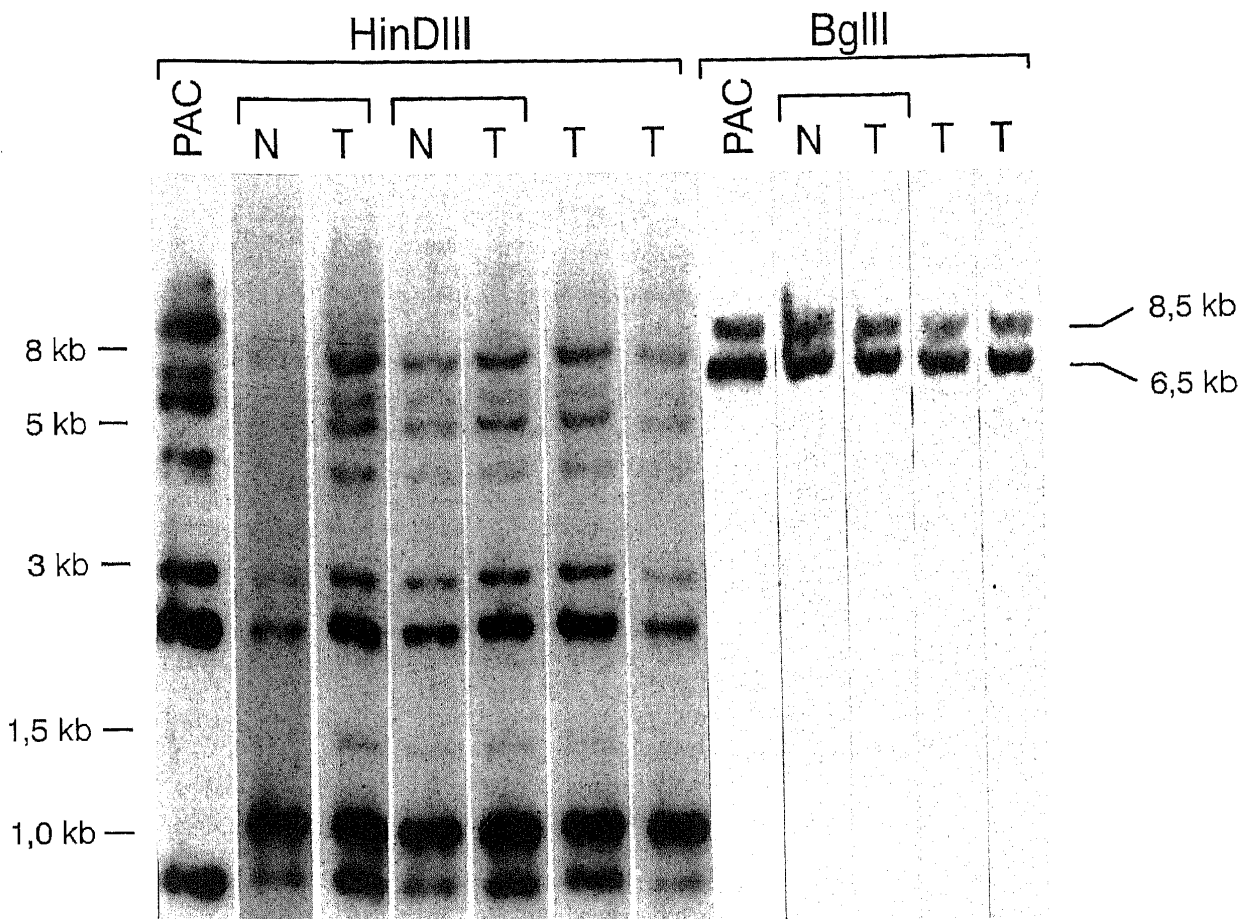


Figure 3. Southern blot analysis of HindIII- and BglIII-digested DNA samples from soft tissue sarcomas and matched normal control tissues and from *TSG101*-specific PAC clone LLNLP704-

P10234Q13. Identical hybridization patterns were obtained in normal tissues (N) and tumor samples (T).

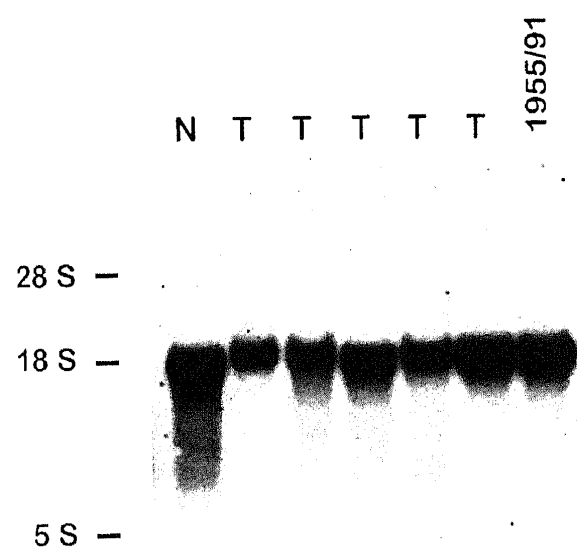


Figure 4. Northern blot analysis of *TSG101* transcription in various soft tissue sarcomas (T), in normal control tissue (N), and in liposarcoma cell line 1955/91. A radiolabeled cDNA probe containing most of the *TSG101* coding region was used for hybridization. All analyzed samples revealed variant transcripts by RT-PCR analysis, but only a single dominant signal is detectable by northern blot hybridization.

frequently both in sarcoma samples and in normal tissues (Figure 2 and Table 2, transcripts A and B). However, all other variant transcripts lacked conserved splice junction sites at the cDNA breakpoints. These findings resemble those from various epithelial malignancies that were reported to show a complex pattern of highly variable, aberrant *TSG101* transcripts lacking both conserved splice donor and acceptor sites and obviously alternatively spliced mRNAs [8,10,11]. The exact mechanism by which the truncated *TSG101* transcripts are generated remains unknown. However, the variability in *TSG101* transcription is further underlined by our observation that different patterns of aberrant transcripts are obtained from recurrent lesions from individual patients (data not shown). For example, variant transcripts were present in a liposarcoma and its second and third metastases, whereas the first metastasis showed only wild-type *TSG101* expression (data not shown). Thus, a relaxed splicing control resulting in the generation of aberrant transcripts is the most likely explanation for these observations.

Both the detection of identical aberrant *TSG101* transcripts in normal tissue and soft tissue sarcoma samples and the very low abundance of truncated *TSG101* transcripts revealed by northern blot hybrid-

ization (Figure 4) do not suggest that *TSG101* variants have major functional relevance in the pathogenesis of soft tissue sarcomas. Furthermore, the observed alterations in the *TSG101* transcripts end up either in frameshift mutations (variant transcripts A, G, H, and K) or in deletions of the postulated ubiquitin-conjugating or coiled coil domains, or both (variant transcripts B, D, and I). The physiological role of each individual truncated *TSG101* gene product has not been analyzed.

In the initial study reporting truncated *TSG101* transcripts in various breast cancers [2], long-distance PCR experiments with genomic *TSG101* sequences suggested the existence of large intragenic deletions that might cause aberrant transcription patterns. In two subsequent studies, however, detailed analyses of the chromosomal *TSG101* locus by Southern blot hybridization both of tumor and of normal tissue samples revealed no large chromosomal rearrangements [8,9]. Similarly, our Southern blot analyses of genomic DNA samples excluded major deletions within the *TSG101* locus in human soft tissue sarcomas. Despite the expression of variant *TSG101* transcripts in the respective tumors, hybridization patterns comparable to those of matched control tissues and to human PAC clones of the *TSG101* locus were obtained (Figure 3).

In conclusion, RT-PCR analysis of soft tissue sarcomas of different histological types revealed the occurrence of aberrant *TSG101* transcripts in more than 60% of the tumor specimens studied. The detection of variant transcripts at comparable frequencies in matched normal tissues and the very low expression of these variants as revealed by northern blot experiments do not support a functional role for variant *TSG101* in the molecular oncogenesis of mesenchymal tumors.

#### ACKNOWLEDGMENTS

We are grateful to Dr. Bernhard Korn (Heidelberg) for performing the high-density filter hybridization of the PAC library and to Dr. Pierre Aman (Lund, Sweden) for kindly providing liposarcoma cell line 1955/91. The graphic support of E. Schierz, J. von Bergmann, and J. Rodrian is greatly appreciated. This work was supported by a grant to FW and CH from

the Deutsche Forschungsgemeinschaft (HE 368/8-1) and by a grant from the Verein zur Förderung der Krebsforschung in Deutschland e.V. to MvKD.

#### REFERENCES

- Li L, Cohen SN. *TSG101* a novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells. *Cell* 85:319-329, 1996.
- Li L, Li X, Francke U, Cohen SN. The *TSG101* tumor susceptibility gene is located in chromosome 11 band p15 and is mutated in human breast cancer. *Cell* 88:143-154, 1997.
- Lothe RA, Hastie N, Heimdal K, Fossa SD, Stenwig AE, Borresen A-L. Frequent loss of 11p13 and 11p15 loci in male germ cell tumours. *Genes Chromosomes Cancer* 7:96-101, 1993.
- Winqvist R, Mannermaa A, Alavaikko M, et al. Refinement of regional loss of heterozygosity for chromosome 11p15.5 in human breast tumors. *Cancer Res* 53:4486-4488, 1993.
- Besnard-Guerin C, Newsham I, Winqvist R, Cavenee WK. A common region of loss of heterozygosity in Wilms' tumor and embryonal rhabdomyosarcoma distal to the D11S988 locus on chromosome 11p15.5. *Hum Genet* 97:163-170, 1996.
- Koonin EV, Abagyan RA. *TSG101* may be the prototype of a class of dominant negative ubiquitin regulators. *Nat Genet* 16:330-331, 1997.
- Ponting CP, Cai Y-D, Bork P. The breast cancer gene product *TSG101*: A regulator of ubiquitination? *J Mol Med* 75:467-469, 1997.
- Lee MP, Feinberg AP. Aberrant splicing but not mutations of *TSG101* in human breast cancer. *Cancer Res* 57:3131-3134, 1997.
- Steiner P, Barnes DM, Harris WH, Weinberg RA. Absence of rearrangements in the tumour susceptibility gene *TSG101* in human breast cancer. *Nat Genet* 16:332-333, 1997.
- Gayther SA, Barski P, Batley SJ, et al. Aberrant splicing of the *TSG101* and *FHIT* genes occurs frequently in multiple malignancies and in normal tissues and mimics alterations previously described in tumours. *Oncogene* 15:2119-2126, 1997.
- Sun Z, Pan J, Buble G, Balk SP. Frequent abnormalities of *TSG101* transcripts in human prostate cancer. *Oncogene* 15:3121-3125, 1997.
- Weiss SW. WHO: Histological Typing of Soft Tissue Tumors. Springer-Verlag, Berlin, 1994.
- Crozat A, Aman P, Mandahl N, Ron D. Fusion of *CHOP* to a novel RNA-binding protein in human myxoid liposarcoma. *Nature* 363:640-644, 1993.
- Willeke F, Ridder R, Mechttersheimer G, et al. Analysis of *FUS-CHOP* fusion transcripts in different types of soft tissue liposarcoma and their diagnostic implications. *Clin Cancer Res* 4:1779-1784, 1998.
- Church GM, Gilbert W. Genomic sequencing. *Proc Natl Acad Sci USA* 81:1991-1995, 1984.
- Sambrook J, Fritsch EF, Maniatis T. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- Ioannou PA, Amemiya CT, Garnes J, et al. A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nat Genet* 6:84-89, 1994.
- Lehrach H, Drmanac R, Hoheisel J, et al. Hybridization fingerprinting in genome mapping and sequencing: Genetic and physical mapping. In: Davies KE, Tilghman SM (eds), *Genome Analysis*, Vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1990, pp 39-81.