Localization of SMAD5 and Its Evaluation As a Candidate Myeloid Tumor Suppressor

Daniel P. Hejlik, Lazar V. Kottickal, Hong Liang, Jeff Fairman, Thaylon Davis, Teresa Janecki, David Sexton, William Perry III, Sean V. Tavtigian, David H. F. Teng, and Lalitha Nagarajan


ABSTRACT

Acquired interstitial or complete losses of chromosome 5 are recurring anomalies associated with preleukemic myelodysplasia and acute myelogenous leukemia with a poor prognosis. Previous studies have delineated a potential myeloid tumor suppressor locus to a <2.4-Mb interval between the genes for IL9 and EGR1 on 5q31. In this report, we have localized the SMAD5 gene, a homologue of the tumor suppressor genes SMAD4/DPC4 and SMAD2/IV18.1, to the minimal myeloid tumor suppressor locus and characterized its open reading frame and genomic organization. SMAD5 transcripts are readily detectable in hematolymphoid tissues and leukemic blasts. Absence of intragenic mutations in the remaining SMAD5 allele of leukemic patients and multiple solid tumor cell lines prescreened for loss of heterozygosity suggests that SMAD5 may not be a common target of somatic inactivation in malignancy.

INTRODUCTION

Acquired loss of all or part of the long arm of human chromosome 5 (5q−) is associated with a wide range of myeloid anomalies (reviewed in Refs. 1 and 2). Identification of a commonly deleted segment at 5q31 has led to the suggestion that loss of gene(s) required for normal myeloid growth and differentiation from the 5q− chromosome unmasks mutations on the remaining allele. This mechanism, originally postulated by Knudson, has been proven in recent years through the biology of many tumor suppressor genes (3).

Invariant LOH5 for several of the polymorphic markers in a 5q31 locus flanked by the genes for IL9 and EGR1 suggests that a critical myeloid tumor suppressor gene may reside in this region (4–6). These findings are consistent with cytogenetic studies that demonstrated that the IL9 gene and the anonymous locus D5S166 flank the smallest region of overlap (7). We have isolated and characterized several single-copy sequences and polymorphic markers from this region and placed these loci with respect to the second-generation genetic map of the human genome (8).

Recent studies have localized the gene for SMAD5 (JV5.1), a homologue of the pancreatic tumor suppressor gene DPC4, to YAC y759D5, which spans the microsatellite markers D5S393 and D5S399 within the IL9-EGR1 interval (9). Members of the SMAD family transduce differentiation and growth arrest signals mediated by the TGF-β family of ligands; loss of function mutations in SMAD genes may confer resistance to the TGF-β family. The amino-terminal DH1/MIH1 and carboxy-terminal DH2/MIH2 domains of SMAD genes are highly homologous among the members and across the species. However, the linker region between these two domains is unique to each SMAD and is conserved in evolution (10). Over seven SMAD homologues have been identified to date; a recent nomenclature4 allows us to distinguish some of the mammalian SMAD genes (11). The genomic localization of SMAD5, coupled with the commonly observed lack of sensitivity of myeloid leukemic cells to TGF-β, makes it an attractive candidate tumor suppressor gene. In this report, we have localized the SMAD5 gene within the physical map of the critical 5q31 locus and evaluated it as a candidate tumor suppressor; a search for intragenic mutations in the remaining allele of leukemic patients and in a panel of malignant cell lines suggests that this gene is unaltered.

MATERIALS AND METHODS

YAC. All of the YACs were obtained from the Human Genome Center at Baylor College of Medicine (Houston, Texas). Coordinates of YACs spanning specific dinucleotide polymorphic loci were obtained from the Whitehead Institute database. The Centre d'Etude du Polymorphisme Humain (CEPH) Mark II library was screened with unique STS markers to isolate additional YACs. Five-ml cultures of colony-purified Saccharomyces cerevisiae cells were grown overnight in yeast peptone dextrose medium. YAC DNA was isolated by the small-scale procedure in which the yeast sporoplasts generated by β-1,3-glucanase treatment were lysed in 0.9% SDS, and the DNA was precipitated by ethanol.

BAC Screens. Oligonucleotides were synthesized for STS within the minimal tiling path. PCR amplification of the STS was performed against the human genomic BAC library (12) from Genome Sciences Inc. (St. Louis, MO). The ends of individual BACs thus isolated were sequenced to develop additional STS for further screening of the BAC library.

cDNAs. The dbEST was searched for cDNAs containing SMAD5 sequences, and the cDNA clones 262686, 206657, 291982, and 214636 thus identified were obtained from the IMAGE consortium (13). Double-stranded sequencing was performed on the largest clone 262686 with multiple oligonucleotide primers. Homology searches were performed by blast analysis of the full-length sequence. Conserved domains were identified with the PILEUP program of the GCG sequence analysis package.

Genomic Sequencing. DNA from BAC clone b37i16 was sequenced with multiple primers derived from the full-length cDNA sequence of clone 262686. Exon-intron boundaries were identified by consensus splice donor and acceptor sites. The SMAD5 cDNA and genomic sequences reported here can be obtained from GenBank (accession numbers AF01601 through AF01607).

Cases and Cytogenetics. Routine cytogenetic analyses were done on all of the MDS/AML patients seen at the Department of Hematology, M. D. Anderson Cancer Center, as detailed elsewhere (4). Samples described in this study were from patients who had either the 5q− chromosome or monosomy for chromosome 5.

Blood Separation, DNA Isolation for Mutation Screens, and LOH Analysis. Peripheral blood specimens collected from MDS and AML patients were separated on Ficoll-Hypaque gradients to yield pelleted granulocyte and buoyant mononuclear leukemia populations. Whenever the sample size per

4 Smad1-JV4.1, Dwarfin-A; Smad2-JV18.1, Madr2; Smad4-DPC4; Smad5-JV5.1, Dwarfin-C.

Received 4/10/97; accepted 7/7/97.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Supported in part by NIH Grants CA66982 and CA55164 and Cancer Center Core Grant CA17662 to the M. D. Anderson Cancer Center. D. P. H. and L. V. K. contributed equally to this work.

1 To whom requests for reprints should be addressed, at Molecular Hematology and Therapy Section; University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Box 81, Houston, Texas 77030. Phone: (713) 792-8756. Fax: (713) 794-5868. E-mail: Lalitha@odin.mdacc.tmc.edu.

2 The abbreviations used are: LOH, loss of heterozygosity; AML, acute myelogenous leukemia; BAC, bacterial artificial chromosome; DH1/MIH1, Drosophila homology domain 1 or MAD homology domain 1; DH2/MIH2, Drosophila homology domain 2 or MAD homology domain 2; MDS, myelodysplasia; ORF, open reading frame; STS, sequence-tagged site(s); TGF-β, transforming growth factor β; YAC, yeast artificial chromosome; IMAGE; integrated molecular analysis of genes and their expression; dbEST, database of expressed sequence tags.
mitted, short-term T-cell cultures were grown from the mononuclear fraction. The cells were lysed, and DNA was isolated with Qiagen columns. Microsatellite analysis was performed as reported previously (5). LOH was determined by the presence of a single allele in the leukemic blast or granulocytic compartment, whereas the lymphoid population revealed two alleles.

**RNA Isolation and Northern Blotting Analysis.** Total RNA was prepared from exponentially growing cells or peripheral blood samples, separated on agarose gels, blotted onto Zeta probe membrane (Bio-Rad), and hybridized with cDNAs labeled with [γ-32P]dCTP by random priming. A Northern blot containing polyadenylated RNA for hematopoietic tissues was purchased from Clonetech Laboratories and probed with both cDNA clones 262866 and 214636 according to the recommended conditions.

**Mutation Screens.** Leukemic, dysplastic samples and solid tumor cell lines that were hemizygous for 5q31 loci were searched for intragenic alterations. Mutation screening was done by PCR cycle sequencing. Outer and nested oligonucleotides were synthesized in the intronic sequences flanking each exon. Because exons 1 and 6 are rather large, additional nested primers were designed within the exon to divide each of these exons into two amplicons. Automated DNA sequencing was facilitated by primer sequences added onto the 5' end of each nested primer. The sequences of these primers are shown in Table 1.

**Table 1. Search for intragenic mutations in leukemic patients and malignant cell lines**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exon 1</td>
<td>AGGAATCGAGGTGGTCAAGAAGAC</td>
</tr>
<tr>
<td>Exon 2</td>
<td>TTCTGATAGCAGGATGAGATC</td>
</tr>
<tr>
<td>Exon 3</td>
<td>TAATCGGTTTCTCAATCGTTAG</td>
</tr>
<tr>
<td>Exon 4</td>
<td>ATTACCAGTCGAGGAGCAGGAGT</td>
</tr>
<tr>
<td>Exon 5</td>
<td>GATATTCTTCTGCAAGCAGGAGG</td>
</tr>
<tr>
<td>Exon 6</td>
<td>TGAATCTAATACTGCTCTTTAAGG</td>
</tr>
</tbody>
</table>

**RESULTS**

**Physical Linkage between the DSS393 Locus and SMAD5 Gene.** We have previously described a YAC contig spanning the myeloid tumor suppressor locus at 5q31. The contig connects the I9L and EGR1 genes and spans the markers DSS393, DSS589, and DSS399 (8). Localization of the SMAD5 gene to the YAC y759D5 containing the loci DSS399 and DSS393 prompted us to fine-map this gene within the minimal tiling path (9). As shown in Fig. 1A, YAC y745F8 with an insert size of 1.3 Mb tested positive for the I9L, DSS5816, SMAD5, DSS393, DSS589, and DSS399 loci, establishing a physical linkage between the I9L and SMAD5 genes. The presence of the I9L and DSS5816 loci in the YAC y11BE8 and DSS5816, SMAD5, DSS393, DSS589, and DSS399 loci in the YAC y880G9 suggests that the SMAD5 gene is telomeric of the I9L gene. Absence of SMAD5 and DSS399 loci in yBSG10 suggests two possible orders: I9L-DSS5816-MAD5-DSS393-DSS589-DSS393 or I9L-DSS5816-DSS393-DSS589/DS399/SMAD5.

Generation of a BAC contig from this locus allowed further ordering of the loci and revealed that the DSS589 locus resides between the DSS393 and DSS399 loci (Fig. 1B). A screen of the BACs with unique STS derived from the 5' end and the 3' untranslated region of the SMAD5 ORF demonstrated that the SMAD5 gene is entirely contained on BAC b37116. Taken together, the BAC and YAC sequencing data are consistent with the order centromere-I9L-DSS5816-MAD5-DSS393-telomere.

**Homology between SMAD5 and Related Proteins.** Beginning with a partial sequence of SMAD5 (JVS.1) reported by Riggins et al. (9), recursive searching of dbEST yielded four IMAGE consortium cDNA clones: 1.1-kb 214636, 1.3-kb 291982, 1.6-kb 206675, and 2.1-kb 262866. Sequence analyses of these four cDNA clones revealed that 262866 encompassed all three shorter clones and that its deduced ORF encoded a 465-amino-acid protein. The DH1/MH1 and DH2/MH2 domains, hallmarks of the SMAD family of proteins, are contained within the amino-terminal 151 residues and the carboxyl-terminal 204 residues, respectively.

As shown in Fig. 2, a search of the GenBank/European Molecular Biology Laboratory protein and DNA databases revealed highest identity (99%) in the deduced ORF to the recently isolated murine Dwarfin C (14). The identity in the DH1/MH1 and DH2/MH2 domains is 100%, and only four conservative differences are seen in the linker region. The most similar Drosophila melanogaster and Caenorhabditis elegans homologues were MAD (80.3% overall identity) and SMA-2 (66.1% overall identity), respectively. Once again, these three genes shared considerable sequence identity in the DH1/MH1 and DH2/MH2 domains (15, 16).

In contrast, homology with other tumor suppressor genes of the SMAD family such as SMAD4/DPC-4 and SMAD2/JVJ8.1 is seen only in the modular DH1/MH1 (<75% identity to SMAD4 and <70% identity to SMAD2) and DH2/MH2 (<70% identity to SMAD4 and <80% identity to SMAD2) domains. Thus, the SMAD5 gene seems to be the human homologue of murine Dwarfin C.

**SMAD5 is Ubiquitously Expressed.** The pattern of SMAD5 expression in the hematopoietic system was examined by Northern blot analysis. As shown in Fig. 3A, we detected two species of RNA (a major 8.7-kb and a minor 4.2-kb species). The expression is high in...
spleen, lymph node, thymus, and fetal liver (Fig. 3A, Lanes 1–3 and 7). There may be an apparent tissue specificity in the expression of the two mRNAs; although low levels are observed in the bone marrow (Fig. 3A, Lane 6) and peripheral blood leukocytes (Fig. 3A, Lane 5), the former has detectable levels of the 4.2-kb transcript, whereas the latter seems to express only the 8.7-kb mRNA. As seen in Fig. 3A, Lane 4, the appendix, isolated from a case of appendicitis and hence heavily contaminated with the hematopoietic cells involved in inflammatory response, transcribes two other novel species of mRNAs (3.2- and 1.9-kb species). An identical pattern is observed when the Northern blot is hybridized with either the full-length cDNA clone 262686 or the 3' clone 214636. The two major transcripts are likely to result from alternate polyadenylation signals in exon 6; furthermore, the 8.7- and 4.2-kb species of mRNA were detected with both probes in multiple tissue Northern blots containing polyadenylated RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas.5

The possibility of alternate polyadenylation signal is further supported by identification of similar-size transcripts with murine Dwarfin C as opposed to a single species of mRNA detected with SMAD5/Dwarfin A (14).

The expression of SMAD5 was next investigated in two AML patients who were hemizygous for this gene. A limited amount of material from these cases restricted our search for expression of the not-so-abundant SMAD5 transcript in total RNA. Nonetheless, the results suggest that, analogous to the normal peripheral blood leukocytes (Fig. 3A, Lane 5), leukemic blasts from patients 12 and 13 express only the 8.7-kb SMAD5 transcript (Fig. 3B, Lanes 1 and 2).

Search for Somatic Mutations in AML and MDS Patients and Tumor Cell Lines. Genomic sequencing of the BAC b37i16 led to the identification of six exons and five introns encoding the entire SMADS ORF. The DH1/MIH1 domain is encoded by exon 1 (138 amino acids) and the first few residues of exon 2. The rest of exon 2 and exon 3 code for the 204-amino acid DH2/MH2 domain, and the polyadenylation site has not yet been identified.

The SMAD5 genomic locus was searched for intragenic mutations in a panel of 27 patients with LOH for 5q31 loci. The results of the mutation screening are summarized in Table 2. No intragenic mutations were detected, although polymorphic single-base alterations could be readily identified under the screening conditions used. To determine whether SMAD5 mutations were associated with malignant transformation of solid tumors, a panel of 45 tumor cell lines prem-screened for LOH in the 5q31 region was examined for intragenic mutations. No gross deletions or inactivating mutations were identified in cell lines derived from malignancies of the pancreas, colon, lung, skin, breast, ovary, and brain (Table 2). Thus, the SMAD5 gene does not seem to be a common target of somatic inactivation in either hematopoietic malignancies or solid tumors.

DISCUSSION

Localization of the SMAD5 gene to a consistent region of loss seen in AML and MDS suggests that it is a candidate myeloid tumor suppressor. Whereas the YAC and BAC contigs presented here are in overall agreement with those in the Whitehead Institute and Lawrence Berkeley National Laboratory databases, our results have further localized the SMAD5 gene between the D5S816 and D5S393 loci within the minimal tiling path.

Chromosome 5 anomalies signify poor prognosis in myeloid malignancies (17). The consistent loss of chromosome 5q31 loci in AML and MDS has suggested the existence of a tumor suppressor gene whose loss confers chemorefractoriness. The DPC-4/SMAD4 gene was originally isolated due to its inactivation in pancreatic tumors, a form of malignancy with poor prognosis (18). Thus, the common clinical outcome between these two states of neoplasia raises the possibility of similar genetic alterations.

The SMAD family genes encode cytoplasmic proteins that are phosphorylated on ligand stimulation to initiate TGF-β-induced growth arrest. The ligand that regulates SMAD5 phosphorylation is

---

5 H. Liang and L. Nagarajan, unpublished observations.
unknown at present, although the murine homologue DWF-C seems to be phosphorylated in response to TGF-\(\beta\) (14). Furthermore, overexpression of murine SMAD5 results in growth inhibition of L6 murine cells upon transient transfections (14). The high level of identity (99%) between the murine and human homologues implies a highly conserved functional role for SMAD5.

An examination of the evolutionary relationship between the structure of the SMA2 gene of \(C.\) \(elegans\) and \(SMAD5\) reveals extensive divergence, despite significant sequence similarity between the two proteins (Fig. 3). In contrast, detection of two mRNA species of similar sizes in the human and mouse distinguishes \(SMAD5\)/\(Dwarfin\) C homologue from other closely related \(SMAD\) family members such as \(Dwarfin\) A/\(SMAD1\) (this report and Ref. 14). Thus, the results presented here serve as a first step toward the characterization of other \(SMAD\) family members to elucidate evolutionary mechanisms responsible for the emergence of the multiple \(SMAD\) genes in mammals.

Two of the \(SMAD\) family members reported to date harbor inactivating mutations in malignant cells; the \(SMAD4/DPC4\) gene is mutated in 50% of pancreatic tumors with allele loss for 18q, whereas the \(SMAD2/JVI\) (this report) and Ref. 14). Thus, the results presented here serve as a first step toward the characterization of other \(SMAD\) family members to elucidate evolutionary mechanisms responsible for the emergence of the multiple \(SMAD\) genes in mammals.

Two of the \(SMAD\) family members reported to date harbor inactivating mutations in malignant cells; the \(SMAD4/DPC4\) gene is mutated in 50% of pancreatic tumors with allele loss for 18q, whereas the \(SMAD2/JVI\) (this report) and Ref. 14). Thus, the results presented here serve as a first step toward the characterization of other \(SMAD\) family members to elucidate evolutionary mechanisms responsible for the emergence of the multiple \(SMAD\) genes in mammals.

Two of the \(SMAD\) family members reported to date harbor inactivating mutations in malignant cells; the \(SMAD4/DPC4\) gene is mutated in 50% of pancreatic tumors with allele loss for 18q, whereas the \(SMAD2/JVI\) (this report) and Ref. 14). Thus, the results presented here serve as a first step toward the characterization of other \(SMAD\) family members to elucidate evolutionary mechanisms responsible for the emergence of the multiple \(SMAD\) genes in mammals. Two of the \(SMAD\) family members reported to date harbor inactivating mutations in malignant cells; the \(SMAD4/DPC4\) gene is mutated in 50% of pancreatic tumors with allele loss for 18q, whereas the \(SMAD2/JVI\) (this report) and Ref. 14). Thus, the results presented here serve as a first step toward the characterization of other \(SMAD\) family members to elucidate evolutionary mechanisms responsible for the emergence of the multiple \(SMAD\) genes in mammals.
are a number of possible explanations for a lack of SMAD5 mutations in leukemic samples: (a) there is a critical myeloid tumor suppressor elsewhere within 5q31; (b) haploinsufficiency may result in a phenotypic alteration; and, (c) there are regulatory mutations that could not be detected under our assay conditions. We favor the first two possibilities because we have not encountered homozygous deletions6 or inactivating mutations for the SMAD5 gene. The possibility of regulatory mutations elsewhere is highly unlikely but cannot be excluded. As for the second possibility, hemizygosity of the D. melanogaster MAD gene results in a phenotype that is halfway between the wild type and the null mutation (15). In addition, hemizygosity of the murine ithub gene, whose D. melanogaster homologue shows a dosage effect, results in expansion of undifferentiated precursors as detected by enhanced expression of CD34, c-kit, RAG-1, and Ikaros genes in embryonic stem cells induced to differentiate along the hematopoietic lineage (20). Future SMAD5 gene knockouts in murine models or mammalian cell lines should reveal the consequences of haploinsufficiency or complete loss of SMAD5 function.

ACKNOWLEDGMENTS

We thank Mark Skolnick for helpful discussions; Patricia Chestnut for critical reading of the manuscript; Drs. M. Andreff, M. Bern, D. Claxton, E. Estey, C. Koller, and S. Komblau for patient material; and Ritesh Mathur for technical assistance.

REFERENCES


Localization of **SMADS** and Its Evaluation As a Candidate Myeloid Tumor Suppressor

Daniel P. Hejlik, Lazar V. Kottickal, Hong Liang, et al.

_Cancer Res_ 1997;57:3779-3783.

Updated version  Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/57/17/3779.citation

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.