

Topoisomerase I and II Consensus Sequences in a 17-kb Deletion Junction of the *COL4A5* and *COL4A6* Genes and Immunohistochemical Analysis of Esophageal Leiomyomatosis Associated with Alport Syndrome

Yasuyoshi Ueki,¹ Ichiro Naito,² Toshitaka Oohashi,¹ Manabu Sugimoto,¹ Tsugio Seki,¹ Hidekatsu Yoshioka,¹ Yoshikazu Sado,³ Hiroshi Sato,⁴ Takashi Sawai,⁵ Fumiaki Sasaki,⁶ Mitsumasa Matsuoka,⁷ Seiji Fukuda,⁸ and Yoshifumi Ninomiya¹

¹Department of Molecular Biology and Biochemistry, Okayama University Medical School, and Divisions of ²Ultrastructural Biology and ³Immunology, Shigei Medical Research Institute, Okayama, Japan; ⁴Second Department of Internal Medicine and ⁵Department of Pathology, Tohoku University School of Medicine, Sendai, Japan; ⁶First Department of Surgery, Hokkaido University School of Medicine, Sapporo, Japan; and ⁷First Department of Surgery, Kumamoto University School of Medicine, and ⁸Kumamoto Rosai Hospital, Kumamoto, Japan

Summary

Diffuse esophageal leiomyomatosis (DL), a benign smooth-muscle-cell tumor, is characterized by abnormal cell proliferation. DL is sometimes associated with X-linked Alport syndrome (AS), an inherited nephropathy caused by *COL4A5* gene mutations. *COL4A5* is tightly linked, in a head-to-head fashion, to the functionally related and coordinately regulated *COL4A6* gene. No X-linked AS cases are due to *COL4A6* mutations, but all DL/AS cases are always associated with deletions spanning the 5' regions of the *COL4A5*/*COL4A6* cluster. Unlike the *COL4A5* breakpoints, those of *COL4A6* are clustered within intron 2 of the gene. We identified a DL/AS deletion and the first characterization of the breakpoint sequences. We show that a deletion eliminates the first coding exon of *COL4A5* and the first two coding exons of *COL4A6*. The breakpoints share the same sequence, which, in turn, is closely homologous to the consensus sequences of topoisomerases I and II. Additional DNA evidence suggested that the male patient is a somatic mosaic for the mutation. Immunohistochemical analysis using α -chain-specific monoclonal antibodies supported this conclusion, since it revealed the absence of the $\alpha 5$ (IV) and $\alpha 6$ (IV) collagen chains in most but not all of the basement membranes of the smooth-muscle-cell tumor. We also documented a similar segmental staining pattern in the glomerular basement membranes of the patient's kidney. This study is particularly relevant to the understanding of DL pathogenesis and its etiology.

Received February 2, 1997; accepted for publication December 11, 1997; electronically published February 6, 1998.

Address for correspondence and reprints: Dr. Yoshifumi Ninomiya, Department of Molecular Biology and Biochemistry, Okayama University Medical School, 2-5-1, Shikata-cho, Okayama 700-8558, Japan. E-mail: yoshinin@ccews2.cc.okayama-u.ac.jp

© 1998 by The American Society of Human Genetics. All rights reserved. 0002-9297/98/6202-0008\$02.00

Introduction

Diffuse esophageal leiomyomatosis (DL; OMIM 308940 [http://www3.ncbi.nlm.nih.gov:80/htbin-post/Omim/dispim?308940]) is a benign smooth-muscle-cell tumor characterized by diffuse proliferation of esophageal smooth-muscle cells and at times is associated with genital and/or bronchial leiomyomatosis. The DL phenotype also has been associated with Alport syndrome (AS; OMIM 301050 [http://www3.ncbi.nlm.nih.gov:80/htbin-post/Omim/dispim?301050]), in rare sporadic as well as familial cases (Garcia-Torres and Orozco 1993). AS is an inherited nephropathy characterized by progressive hematuria often associated with sensorineural deafness and ocular lesions. The syndrome is inherited usually as an X-linked dominant trait or, less frequently, as an autosomal recessive trait. Since the first discovery of the molecular defect in X-linked AS, >120 mutations have been identified in the *COL4A5* gene (Knebelmann et al. 1996; Renieri et al. 1996). *COL4A5* codes for the common subunit of the type IV collagen trimers in the glomerular basement membrane. Mutations also have been reported in the genes for the $\alpha 3$ (IV) and $\alpha 4$ (IV) chains (*COL4A3* and *COL4A4*, respectively) of individuals with the rare, autosomal recessive form of AS (Lemmink et al. 1994; Mochizuki et al. 1994). In contrast, no AS case with a mutant *COL4A6* gene has ever been reported. Like the *COL4A3*/*COL4A4* pair on chromosome 2, *COL4A5* and *COL4A6* are tightly linked on chromosome X, in a head-to-head organization, and share the same promoter (Oohashi et al. 1994, 1995; Sugimoto et al. 1994).

Molecular studies recently have established a causal relationship between the DL/AS phenotype and genomic deletions extending from the 5'-end portion of *COL4A5* to the 5'-end portion of *COL4A6* and including the bidirectional promoter (Zhou et al. 1993). Unlike the *COL4A5* breakpoints, the positions of which varied

among different patients, another breakpoint was consistently found within a relatively small segment of the ~180-kb intervening sequence (IVS) III of *COL4A6* (Kendall et al. 1997). Interestingly, larger deletions, extending beyond exon 3 of *COL4A6*, did not cause DL and only caused AS (Heidet et al. 1995). The authors therefore suggest that the IVS III deletion leads either to the generation of a truncated *COL4A6* product that somehow is involved in DL pathogenesis or to the loss of a gene within *COL4A6* that, presumably, regulates smooth-muscle-cell proliferation.

As a first step toward unraveling of the mechanism(s) underlying DL, the present study was designed to elucidate the composition of the deletion region of the genomic breakpoints in a DL/AS case. The results implicate the possible involvement of topoisomerase enzymes in the genomic rearrangement responsible for the DL/AS phenotype and raise the possibility that there might be a third gene in IVS III that is involved in the regulation of smooth-muscle-cell proliferation.

Patients, Material, and Methods

Patient

The patient was a 19-year-old boy who, at the age of 6 years, was diagnosed with hematuria. A kidney-biopsy specimen revealed only a mild mesangial proliferation. Samples subjected to electron microscopic analysis contained no glomeruli; therefore, ultrastructural analysis of the glomerular basement membrane could not be performed. At the age of 8 years, esophageal dysfunction manifested itself, and the patient underwent esophago-gastrectomy. The histological diagnosis was DL associated with AS. The patient's renal function had gradually worsened. There is no history of hematuria or esophageal tumors in the patient's family.

White Blood Cells and Southern Blot Analysis

Total genomic DNA was extracted and purified from peripheral white blood cells of the patient and was used for Southern blot analysis. No samples were available from the patient's parents. Some of the peripheral white blood cells of the patient were used to establish Epstein-Barr-virus-derived lymphoblastoid cell lines, in accordance with the study by Maddalena et al. (1988). Southern blot hybridizations were performed by following routine protocols (Sambrook et al. 1989). Likewise, a commercially available EMBL3 human genomic library (HL1111J and HL1006dn; Clontech) was screened, and the resultant positive clones were purified and characterized in accordance with standard procedures (Sambrook et al. 1989).

PCR Amplification

The patient's genomic DNA was amplified with primers SK F2 (5'-ATA TAG CTG GAA AAG TGT GC-3') and YU20E1.8R1 (5'-TAA CTG GGA AAT GTC AGA TAA C-3'). To make the 260-bp probe 2 (fig. 1B), we designed a 5' primer (5'-GGA CAC ACA CCA TGA CAT GG-3') and a 3' primer (5'-ACC TGC GAA ACC ATC ACA AC-3'). The reaction was performed by use of 50 ng genomic DNA, 0.5 mM each primer, 0.8 mM dNTPs, 1.5 mM Mg(OAc)₂, 0.5 U *rTth* polymerase XL, and 1 × XL buffer (Perkin Elmer), in a total volume of 25 μl. The PCR conditions were as follows: denaturation at 94°C for 1 min and 30 cycles of 94°C for 30 s, 58°C for 1 min, and 72°C for 3 min, with a final elongation at 72°C for 10 min. PCR products were subcloned in TA vectors (Invitrogen) and were sequenced with an ABI 373S automatic sequencer (Central Research Laboratory, Okayama University Medical School). Additional oligonucleotides were synthesized for the subcloning and sequencing of the deletion junction.

Immunohistochemical Staining

Normal portions of the esophagus were obtained from an 80-year-old patient with stomach cancer, and normal portions of kidney tissues were obtained from a 70-year-old patient with renal cancer. Esophageal tumor samples from the DL/AS patient were fixed in formaldehyde and were embedded in paraffin. Previously established monoclonal antibodies specific for the six α(IV) chains were used for differential staining (Ninomiya et al. 1995). Each sample was deparaffinized and stained under several conditions, as described below. Since fresh renal sections were not available, paraffin-embedded sections were used for the analysis of type IV collagen chains. Accordingly, the paraffin-embedded sections were subjected to an antigen-retrieval procedure (Shin et al. 1991). In brief, sections were dipped in 0.2 M HCl (pH 0.9) and then were heated in a small autoclave (KT2322; ALP Co., Ltd.), at different temperatures, for 6 min. The suitable temperature for autoclaving was established empirically for each specimen and basement membrane (Naito et al. 1996). Optimal treatment for the kidney sections was at 110°C for 6 min, whereas that for the esophageal tumors was at 121°C or 127°C. After antigen retrieval, the paraffin-embedded sections were incubated with antibodies, at room temperature, for 60 min. After washing with phosphate-buffered saline, the antibodies were detected with a Labeled Streptavidin-biotin kit (DAKO Co.).

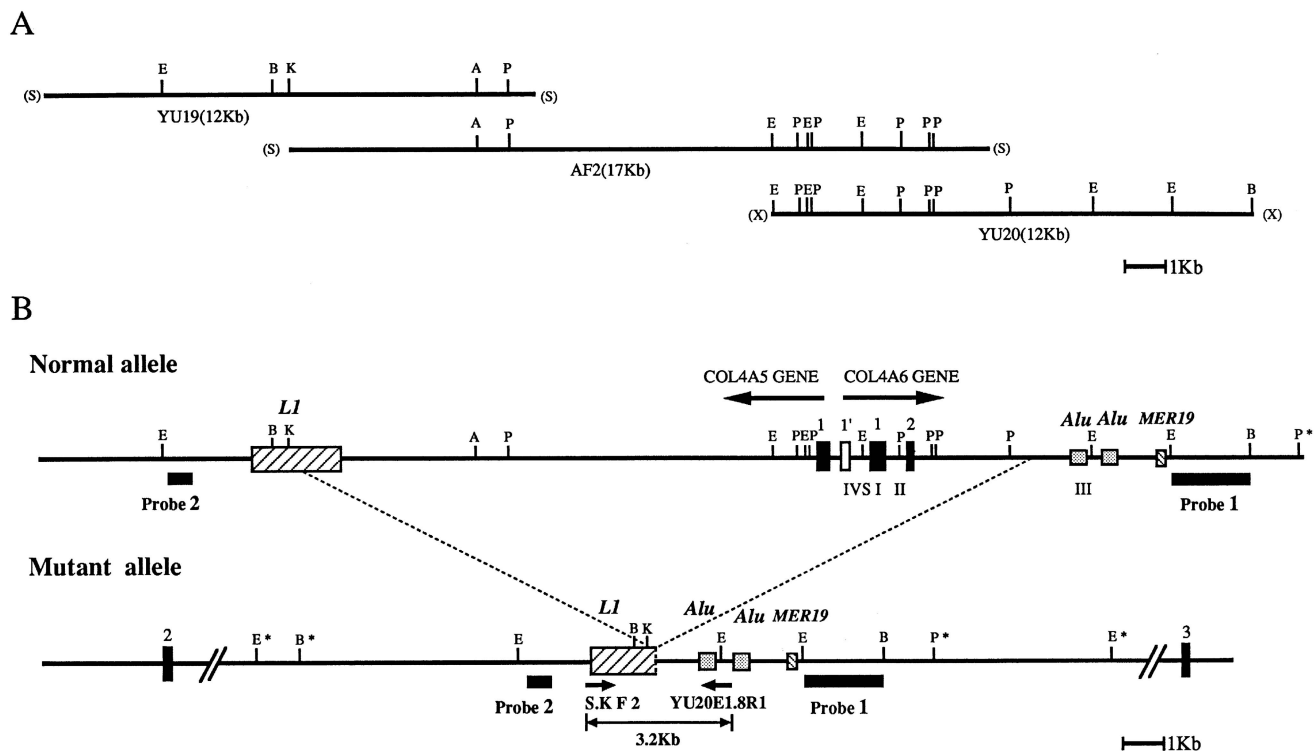


Figure 1 Relative location of the three genomic DNA fragments (A), and restriction map of genomic DNA of the normal allele and the patient’s mutant allele (B). The relative location of the three overlapping clones is shown, and representative restriction sites are indicated by uppercase letters: E = *EcoRI*; S = *SalI*; X = *XhoI*; B = *BamHI*; K = *KpnI*; A = *ApaI*; and P = *PstI*. Uppercase letters within parentheses indicate cloning sites. Uppercase letters with an asterisk (*) indicate the restriction sites that were estimated by genomic Southern hybridization. The name of each clone is given under the clone, and the numbers within parentheses indicate the size of each clone. Restriction sites of the normal allele (B) are drawn in correspondence with the three isolated clones. As indicated, the two genes, *COL4A5* and *COL4A6*, are closely located and are arranged head to head. Exons 1 and 1' of the *COL4A6* gene are used alternatively (Sugimoto et al. 1994). For better understanding, only *COL4A6* introns are indicated, as IVS I, II, and III. Blackened and unblackened boxes indicate exons, with the exon number given above the box. Several characteristic elements (i.e., L1, *Alu*, and the MER19 homologous sequence) are indicated by different types of boxes (Scott et al. 1987; Britten et al. 1988; Kaplan et al. 1991). The location of the 17-kb deletion is indicated on both the normal and the mutant alleles. Probes 1 and 2, used for the genomic Southern hybridization (fig. 2), and the two PCR primers are shown by blackened bars and by the two small arrows, respectively.

Results

Identification of the DL/AS Deletion and Evidence for Mosaicism

Since all DL/AS cases reported to date have a deletion of the first three exons (1', 1, and 2) of *COL4A6*, we isolated overlapping genomic clones from this region, for detailed analysis of our patient. These were obtained after screening a library containing normal human DNA, with the sense *COL4A6* cDNA. After isolation of the first positive clone (AF2; fig. 1A), the screen was repeated with AF2 end probes, resulting in the identification of two additional recombinants (YU19 and YU20; fig. 1A). Together, they extended for >30 kb of DNA, from IVS I of *COL4A5* to IVS III of *COL4A6* (fig. 1B).

A 1.8-kb genomic probe corresponding to the 5' segment of IVS III of *COL4A6* (probe 1; fig. 1B) was used

in the Southern analysis of the patient’s DNA digested with *BamHI*. On the basis of the restriction map of the genomic clones, the expected size of the normal allele was 23 kb (fig. 1B). Consistent with this prediction, a single band of 23 kb was seen in the control samples, prepared from an unaffected male (fig. 2A, lane 2) and an unaffected female (fig. 2A, lane 3), with the intensity of the former noticeably weaker than that of the latter. In contrast, a new and smaller band of ~6 kb was noted for the patient’s DNA (fig. 2A, lane 1). The opposite size differences were observed with *PstI*-digested DNA hybridized to the same probe (data not shown). When considered together, the results suggested a 17-kb deletion extending from 3' of the third *PstI* site of IVS III of *COL4A6* to immediately 5' of the *BamHI* site of IVS I of *COL4A5* (fig. 1B).

Interestingly, both the *BamHI* (fig. 2A) and *PstI* (data

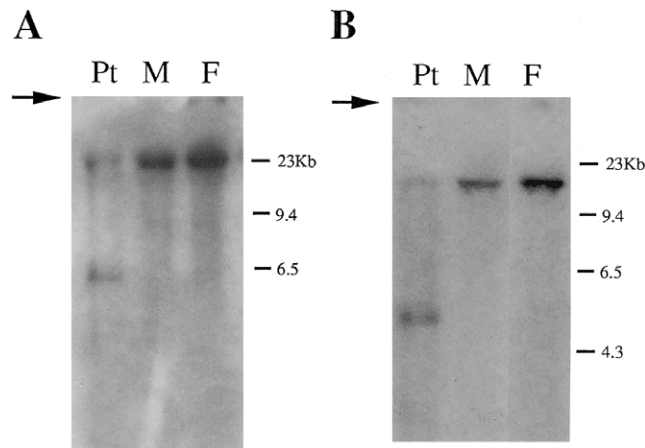


Figure 2 Results of genomic Southern hybridization. Total genomic DNA was digested with (A) *Bam*HI or (B) *Eco*RI. The blots were hybridized with (A) probe 1 (fig. 1B) or (B) probe 2, located in *COL4A5* intron 1 (fig. 1B). Five micrograms of DNA were loaded on each lane. Lane Pt, DNA extracted from peripheral white blood cells of the DL/AS patient. Lane M, Unaffected male control. Lane F, Unaffected female control. Note that the 6-kb band appears in the patient's DNA (A).

not shown) digests of the DL/AS sample also displayed the bands corresponding to the normal allele. The identical finding was noted when *Eco*RI-digested DNA from the patient's peripheral blood was hybridized to a probe corresponding to the 5' portion of IVS I of *COL4A5* (probe 2; fig. 2B). However, in all samples, the so-called mutant-allele band was consistently darker than the so-called normal-allele band, thus raising the possibility of somatic mosaicism.

Characterization of the Deletion Breakpoints

The results described above suggest that the DL/AS mutation is within the 23-kb *Bam*HI fragment of the normal allele (fig. 1B). Thus, we next focused on determining the nucleotide sequences around the presumptive breakpoints, one around the *Bam*HI site near probe 2 and the other around the *Pst*I and *Eco*RI sites near probe 1. To determine these sequences, two primers were designed to verify the deletion, by PCR amplification of genomic DNA. Consistent with our prediction that only the mutant allele would amplify, only the mutant 3.2-kb PCR product was obtained after amplification of DNA from the patient's peripheral blood (fig. 3). This product was sequenced in its entirety and was compared with the respective segments of the *COL4A5* and *COL4A6* genes.

The comparison demonstrated that the *COL4A5* side of the breakpoint resides within the L1 transposable element (as mentioned below) and that the *COL4A6* side is at nucleotides 3570-3576 of IVS III of the *COL4A6*

gene. Therefore, the deletion includes 12 kb of intron 1, exon 1 of *COL4A5*, the bidirectional promoter, and from exon 1' to 3 kb within IVS III of *COL4A6* (see fig. 1B). Nucleotide-sequence alignment around the breakpoints showed that they display CATCATC junctional homology (fig. 4). Interestingly, the breakpoints have six topoisomerase I consensus sequences on both strands. The *COL4A6* breakpoint also exhibits a 90% match to the topoisomerase II consensus sequence. A 6-bp direct repeat, AAATGC, and a sequence homologous to *Alu* also were located near both the *COL4A5* and *COL4A6* breakpoints. There was no nucleotide insertion or deletion in the junctional fragment.

Mosaic Loss of Normal Basement-Membrane Architecture

In order to relate the genotypic findings to the patient's phenotype, we immunohistochemically analyzed the basement membranes surrounding the smooth-muscle cells at the lower esophageal region and those of kidney glomeruli, using antibodies for each of the type IV collagen chains. The esophageal sample from an unaffected individual showed positive staining with monoclonal antibodies specific for the α 1(IV), α 2(IV), α 5(IV), and α 6(IV) chains (fig. 5A, B, E, and F, respectively). As was expected, antibodies against the α 3(IV) and α 4(IV) chains did not show any positive signals in this tissue (fig. 5C and D, respectively). In the DL/AS sample, the same antibodies visualized the α 1(IV) and α 2(IV) chains (fig. 5G and H, respectively) and failed to do so with

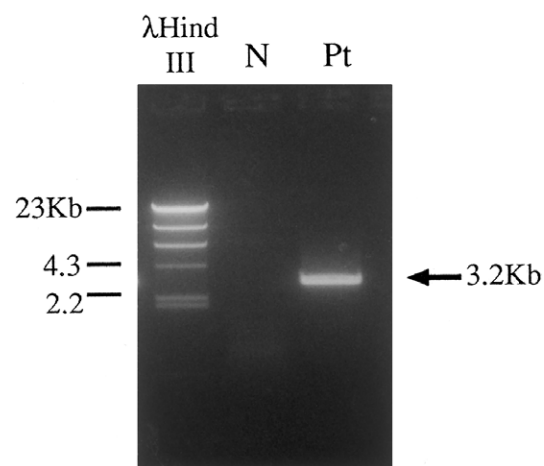


Figure 3 Results of PCR amplification of the junctional fragment from DNA of the DL/AS patient. Lane 1, λ *Hind*III size markers. Lane 2, Normal control DNA used as a template. Lane 3, DL/AS patient's DNA from peripheral white blood cells used as a template. Note that the amplified band from the two primers, SK F2 and YU20E1.8R1, was 3.2 kb. This fragment was subcloned, and the nucleotide sequence was determined. The relative locations of the primers and the amplified fragment are indicated in figure 1B.

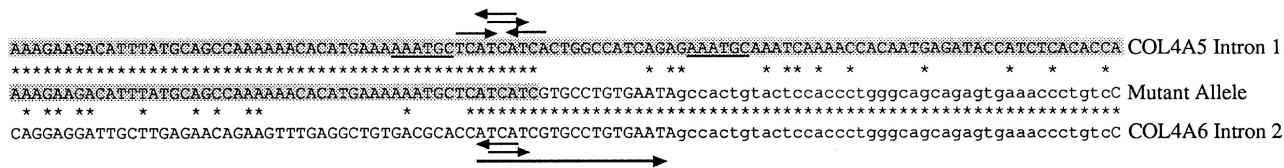


Figure 4 Nucleotide sequences of normal and mutant alleles, at the breakpoints. The relative location of the 17-kb deletion is shown in figure 1B. Nucleotide sequences of part of intron 1 of *COL4A5* (top), part of IVS III of *COL4A6* of the normal allele, across the breakpoints (bottom), and the corresponding region from the mutant allele in one strand (middle) are aligned. The asterisks (*) indicate the nucleotide matches. Note that the *COL4A5* intron 1 sequence and the mutant-allele sequence match completely up to the central CATCATC sequence. The 3' side of the CATCATC sequence of the mutant allele and the *COL4A6* IVS III sequence are also the same. Short and long arrows indicate the consensus sequences of topoisomerase I (5'-A/T-G/C-T/A-T-3') (Been et al. 1984) and topoisomerase II (5'-A/G-N-T/C-NNCNNG-T/C-NG-G/T-TN-T/C-N-T/C-3') (Spitzner and Muller 1988), respectively. Underlined sequences indicate the 6-bp direct repeat. Shaded letters and lowercase letters represent the L1 sequence and the sequence homologous to *Alu*, respectively.

the $\alpha 3(IV)$ and $\alpha 4(IV)$ chains (fig. 5I and J, respectively). However, the expression patterns of the $\alpha 5(IV)$ and $\alpha 6(IV)$ chains were significantly different from those of the control samples. With the exception of a few smooth-muscle cells, most of the basement membranes were unreactive with these two antibodies (fig. 5K and L). Aside from differing from the normal pattern, the uneven distribution of the $\alpha 5$ and $\alpha 6$ epitopes in the DL/AS sample indirectly supports the DNA evidence for somatic mosaicism of the *COL4A5*/*COL4A6* deletion.

In a normal kidney-cortex sample, the $\alpha 1(IV)$ and $\alpha 2(IV)$ chains in all basement membranes were stained (fig. 5M and N, respectively). Likewise, the $\alpha 3(IV)$, $\alpha 4(IV)$, and $\alpha 5(IV)$ chains showed the same positive expression pattern in glomerular basement membranes and in some parts of the tubular basement membranes (fig. 5O, P, and Q, respectively). Consistent with our previous data, anti- $\alpha 6(IV)$ did not show any positive staining with glomerular basement membranes but did so with some tubular and the Bowman basement membranes (fig. 5R) (Ninomiya et al. 1995). In the patient's sample, the expression patterns of the $\alpha 1(IV)$ and $\alpha 2(IV)$ chains were the same as those in normal tissue (fig. 5S and T, respectively). However, the staining patterns of the $\alpha 3(IV)$, $\alpha 4(IV)$, and $\alpha 5(IV)$ chains in glomerular basement membranes were segmental (fig. 5U, V, and W, respectively). A similar discontinuous pattern was observed for the $\alpha 6(IV)$ antibody, in the Bowman and in some tubular basement membranes (fig. 5X). Hence, the affected tissues of the DL/AS patient displayed a pathology consistent with the loss of *COL4A5* and *COL4A6* gene products in most, but not all, the cells.

Discussion

Deletions are a common type of mutation found in many organisms. The characterization of a deletion consisting of the juxtaposition of noncontiguous sequences as a result of the breaking and rejoining of the DNA

strands provides an excellent opportunity to study the mechanisms that are involved in the generation of mutations in the mammalian genome. A previous report involved the partial characterization of the deletions for DL/AS (Heidet et al. 1995). However, in this paper, we have reported the first detailed analysis of a 17-kb deletion that included the upstream regions of both the *COL4A5* and *COL4A6* genes and the presumptive common promoter region for both genes, from a case of DL/AS. The deletion encompassing the two collagen genes caused two disease states, DL and AS, although one of the genes, *COL4A5*, is already known to be linked to AS.

It has been postulated that topoisomerases I and II may be involved in recombinational events, since these enzymes form covalent linkages in both strands of the DNA helix and can cause a single-strand break and a transient double-strand break in the helix (Bae et al. 1988; Champoux and Bullock 1988). Hu et al. (1991) suggested that the enzymes are able to cleave DNA strands, during duplication of the dystrophin gene, and several investigations have reported that topoisomerase I and II sites were found at the breakpoints of translocations or deletions in the dystrophin and steroid sulfatase genes (Shapiro et al. 1989; Bodrug et al. 1991; van Bakel et al. 1995). In the DL/AS case, the consensus sequences of topoisomerase I and II binding sites were found at the breakpoints. In particular, six topoisomerase I consensus sequences existed on both strands, in the CATCATC junctional homologous sequence. These results suggest that topoisomerase I and II activity may be involved in the mechanisms of DNA breakage and reunion.

The common mechanism involved in large gene deletions is recombination. There are two types of recombination events: one is the homologous unequal recombination between gene sequences or between repetitive sequence elements, and the other is nonhomologous recombination between DNA sequences with short ho-

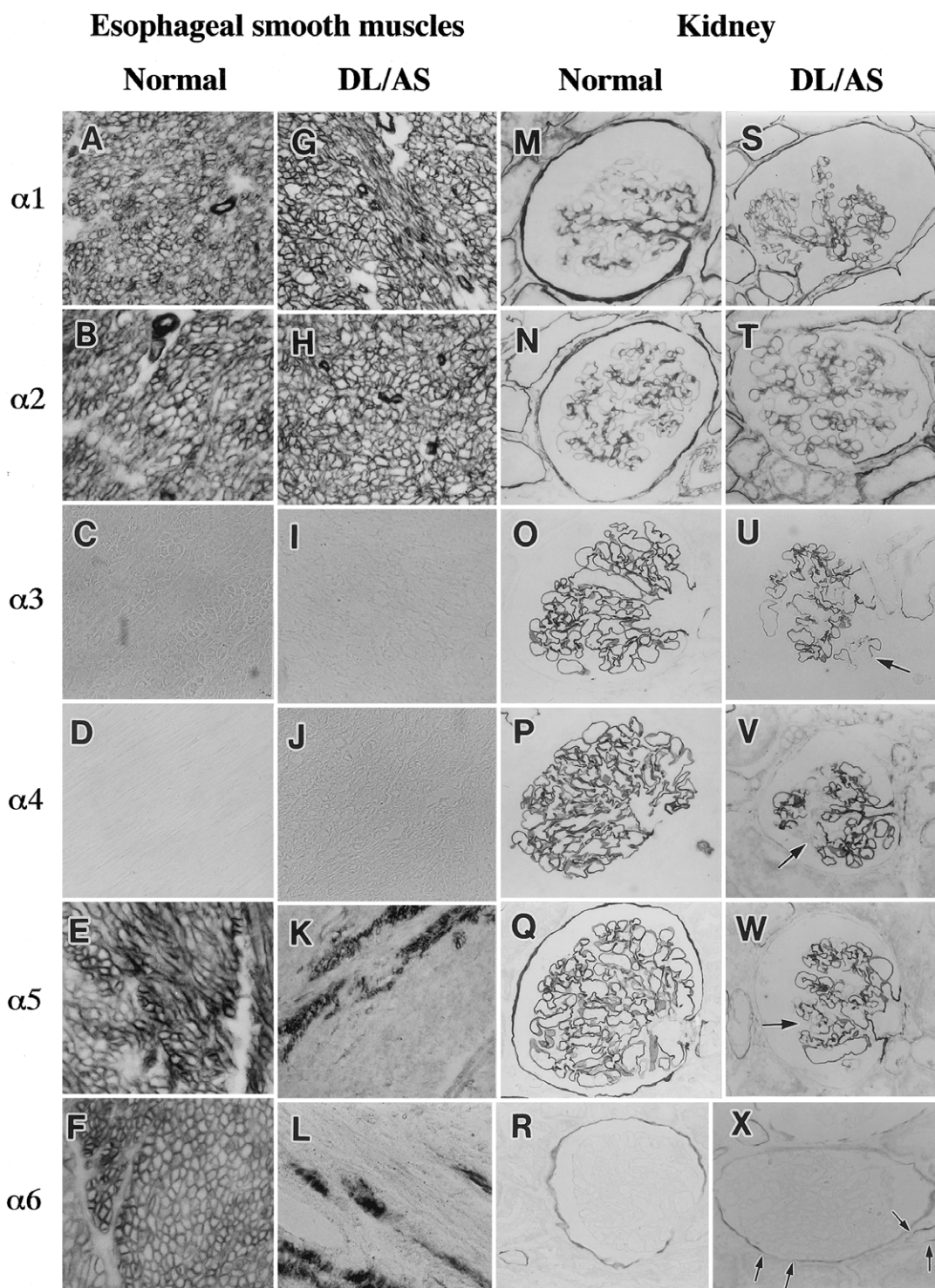


Figure 5 Distribution of $\alpha(IV)$ chains in esophageal smooth muscles and kidney from a normal control and from the DL/AS patient. $\alpha(IV)$ -Chain-specific monoclonal antibodies (H11, H21, H31, H44, H52, and H63) (Ninomiya et al. 1995) were used for immunostaining. When smooth muscles in the lower esophageal region were stained, antibodies against $\alpha 1$, $\alpha 2$, $\alpha 5$, and $\alpha 6$ chains (A, B, E, and F, respectively), but not those against $\alpha 3$ and $\alpha 4$ chains (C and D, respectively), showed positive signals in the basement membranes surrounding the smooth-muscle cells. However, although the expression patterns for $\alpha 1$, $\alpha 2$, $\alpha 3$, and $\alpha 4$ chains were almost the same in tumor tissues (G, H, I, and J, respectively) as in control tissues (A, B, C, and D, respectively), most of the basement membranes around smooth-muscle cells were negative for $\alpha 5$ and $\alpha 6$ antibodies (K and L, respectively). In the normal kidney, $\alpha 1$ and $\alpha 2$ chains were coexpressed in all basement membranes (M and N, respectively), whereas $\alpha 3$, $\alpha 4$, and $\alpha 5$ chains were colocalized in glomerular basement membranes (O, P, and Q, respectively), and $\alpha 5$ and $\alpha 6$ chains were coexpressed in Bowman basement membranes (Q and R, respectively). Distribution of $\alpha 1$ and $\alpha 2$ chains was the same in the patient's kidney (S and T, respectively) as in that of the control (M and N, respectively). However, $\alpha 3$, $\alpha 4$, and $\alpha 5$ chains showed a segmental pattern (indicated by arrows) in glomerular basement membranes (U, V, and W, respectively). In Bowman capsule and tubular basement membranes, the $\alpha 6$ chain showed a discontinuous pattern (X, between the two small arrows).

mology (Cooper and Krawczak 1993). Efstratiadis et al. (1980) reported that, in eukaryotic genes, short direct repeats are involved in the generation of deletions by slipped mispairing during DNA replication. Nonhomologous recombination could cause large DNA rearrangements (Albertini et al. 1982). In the β -globin gene of HPFH-2 (*hereditary persistence of fetal hemoglobin*), a short direct repeat was observed at the junctions of a >100-kb deletion (Henthorn et al. 1990). The deletion of the *COL4A5* and *COL4A6* genes presented here could have involved the short repeat sequence CAT-CATC, located only at both breakpoints. Nicked or gapped breakpoints caused by topoisomerase I and II could contribute to a single-strand loop excision during slipped mispairing.

Another interesting finding was that the breakpoint on the *COL4A5* side was located within the L1 repetitive element. Several examples of large deletions at the L1 elements have been reported, including that in the β -globin gene (Henthorn et al. 1990), although many more deletions through *Alu* repetitive elements have occurred in the genome (Ottolenghi and Giglioni 1982; Lehrman et al. 1985, 1986; Li and Bray 1993).

Genomic Southern analysis using DNA from peripheral white blood cells suggested somatic mosaicism. Immunohistochemical data from the basement membranes of glomeruli and of some tubules also supported this idea. Interestingly, it is the smooth-muscle cells proliferating preferentially in the esophageal region that cause the DL phenotype. This could be because the mutant cells have gained a growth abnormality, owing to the 17-kb deletion at the Xq22 region. We think that this implies the possibility that a large deletion in the *COL4A6* gene, during somatic mitosis, could cause DL, even if the lesion is limited within the esophagus.

The expression of the *COL4A5* and *COL4A6* genes was blocked completely in most of the tumor tissues. Heidet et al. (1995, 1997) had detected transcription from exon 4 of the *COL4A6* gene, by reverse-transcriptase PCR, and had suggested the possibility that an abnormal truncated $\alpha 6(\text{IV})$ chain could cause DL. However, we think it unlikely that the triple-helix collagen molecules that contain the abnormal truncated $\alpha 6(\text{IV})$ chains could be responsible for the pathogenesis of DL, because we could not detect, by immunohistochemical analysis, collagen molecules containing the $\alpha 6(\text{IV})$ chain, in most of the smooth-muscle cells of the tumor. What, then, causes DL in the cases of deletion of the upstream region of both the *COL4A5* and *COL4A6* genes? Since one case having a large deletion containing the entire *COL4A6* gene represented only the AS and not the DL phenotype (Heidet et al. 1995), it is not plausible that the entire loss of *COL4A6* expression causes DL.

Our hypothesis is that a third gene and/or its regulatory elements, within the 180-kb IVS III (Kendall et

al. 1997), may regulate smooth-muscle-cell proliferation. In our DL/AS patient, the 17-kb deletion may have hit one of the regulatory elements and thus may have changed the expression of the third gene. Several AS cases that contained larger deletions, ranging from *COL4A5* IVS I to *COL4A6* IVS IV, did not show the DL phenotype (Heidet et al. 1995). We think that the deletions in these AS cases all include the entire *COL4A6* IVS III, which presumably contains the third gene, together with its regulatory elements, and therefore no obvious phenotypic changes were detected. A similar situation has been reported for other benign mesenchymal tumors, such as uterine leiomyoma, lipoma, and hamartoma (Schoenmakers et al. 1995). Schoenmakers et al. demonstrated that a gene for one of the high-mobility-group proteins, *HMGI-C*, was disrupted, and fusion transcripts were detected in several benign solid tumors, suggesting a link between a gene for a member of the high-mobility-group proteins and development of benign solid tumors. Other studies, reported by de Kok et al. (1995, 1996), have suggested that X-linked deafness can be caused by an 8-kb deletion ~900 kb from the *POU3F4* gene. de Kok et al. (1995, 1996) assumed that the 8-kb region contains some *cis* regulatory elements (a promoter and/or an enhancer) for transcription of the gene. We also believe that the 17-kb deleted region, especially the *COL4A6* IVS III, contains several essential genes and/or *cis* elements that could regulate smooth-muscle-cell proliferation, as mentioned above. As an approach to answer this question, gene targeting using constructs lacking the entire or a part of IVS III could be performed. However, the possibility that the truncated $\alpha 6(\text{IV})$ chain could cause the DL phenotype cannot be ignored (Heidet et al. 1995). This possibility can be tested by experiments on cDNA transfection of the truncated form of the $\alpha 6(\text{IV})$ chain into smooth-muscle cells.

Acknowledgments

We would like to thank our patients for their participation in the study. We also thank Drs. Tomu Hayashi and Sumio Hasegawa, Tohoku University School of Medicine, for providing samples and for their stimulating discussion. This work was supported by Grants-in-Aid for Scientific Research and by research grants from the Yamanouchi Foundation for Research on Metabolic Disorders and from the Ryobi Teien Foundation. The nucleotide sequence data reported in this paper have been deposited in the DNA Data Bank of Japan (<http://www.DDBJ.nig.ac.jp/>), under accession number AB008499.

References

Albertini AM, Hofer M, Calos MP, Miller JH (1982) On the formation of spontaneous deletions: the importance of short

- sequence homologies in the generation of large deletions. *Cell* 29:319-328
- Bae YS, Kawasaki I, Ikeda H, Liu LF (1988) Illegitimate recombination mediated by calf thymus DNA topoisomerase II *in vitro*. *Proc Natl Acad Sci USA* 85:2076-2080
- Been MD, Burgess RR, Champoux JJ (1984) Nucleotide sequence preference at rat liver and wheat germ type 1 DNA topoisomerase breakage sites in duplex SV40 DNA. *Nucleic Acids Res* 12:3097-3114
- Bodrug SE, Holden JJA, Ray PN, Worton RG (1991) Molecular analysis of X-autosome translocations in females with Duchenne muscular dystrophy. *EMBO J* 10:3931-3939
- Britten RJ, Baron WF, Stout DB, Davidson EH (1988) Sources and evolution of human *Alu* repeated sequences. *Proc Natl Acad Sci USA* 85:4770-4774
- Champoux JJ, Bullock PA (1988) Possible role for the eukaryotic type I topoisomerase in illegitimate recombination. In: Kucherlapati R, Smith GR (eds) *Genetic recombination*. American Society for Microbiology, Washington DC, 655-666
- Cooper DN, Krawczak M (1993) *Human gene mutation*. BIOS Scientific, Oxford
- de Kok YJM, van der Maarel SM, Bitner-Glindzic M, Huber I, Monaco AP, Malcolm S, Pembrey ME, et al (1995) Association between X-linked mixed deafness and mutations in the POU domain gene *POU3F4*. *Science* 267:685-688
- de Kok YJM, Vossenaar ER, Cremers CWRJ, Dahl N, Laporte J, Hu LJ, Lacombe D, et al (1996) Identification of a hot spot for microdeletions in patients with X-linked deafness type 3 (DFN3) 900 kb proximal to the DFN3 gene *POU3F4*. *Hum Mol Genet* 5:1229-1235
- Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, et al (1980) The structure and evolution of the human β -globin gene family. *Cell* 21:653-668
- Garcia-Torres R, Orozco L (1993) Alport-leiomyomatosis syndrome: an update. *Am J Kidney Dis* 22:641-648
- Heidet L, Cai Y, Sado Y, Ninomiya Y, Thorner P, Guicharnaud L, Boye E, et al. (1997) Diffuse leiomyomatosis associated with X-linked Alport syndrome: extracellular matrix study using immunohistochemistry and *in situ* hybridization. *Lab Invest* 76:233-243
- Heidet L, Dahan K, Zhou J, Xu Z, Cochat P, Gould JDM, Leppig KA, et al (1995) Deletions of both $\alpha 5(IV)$ and $\alpha 6(IV)$ collagen genes in Alport syndrome and in Alport syndrome associated with smooth muscle tumors. *Hum Mol Genet* 4: 99-108
- Henthorn PS, Smithies O, Mager DL (1990) Molecular analysis of deletions in the human β -globin gene cluster: deletion junctions and locations of breakpoints. *Genomics* 6: 226-237
- Hu X, Ray PN, Worton RG (1991) Mechanisms of tandem duplication in the Duchenne muscular dystrophy gene include both homologous and nonhomologous intrachromosomal recombination. *EMBO J* 10:2471-2477
- Kaplan DJ, Jurka J, Solus JF, Duncan CH (1991) Medium reiteration frequency repetitive sequences in the human genome. *Nucleic Acids Res* 19:4731-4738
- Kendall E, Evans W, Jin H, Holland J, Vetric D (1997) A complete YAC contig and cosmid interval map covering the entirety of human Xq21.33 to Xq22.3 from DXS3 to DXS287. *Genomics* 43:171-182
- Knebelmann B, Breillat C, Forestier L, Arrondel C, Jacassier D, Giatras I, Drouot L, et al (1996) Spectrum of mutations in the COL4A5 collagen gene in X-linked Alport syndrome. *Am J Hum Genet* 59:1221-1232
- Lehrman MA, Russell DW, Goldstein JL, Brown MS (1986) Exon-*Alu* recombination deletes 5 kilobases from the low density lipoprotein receptor gene, producing a null phenotype in familial hypercholesterolemia. *Proc Natl Acad Sci USA* 83:3679-3683
- Lehrman MA, Schneider WJ, Sudhof TC, Brown MS, Goldstein JL, Russell DW (1985) Mutation in LDL receptor: *Alu-Alu* recombination deletes exons encoding transmembrane and cytoplasmic domains. *Science* 227:140-146
- Lemmink HH, Mochizuki T, van den Heuvel LPWJ, Schroder CH, Barrientos A, Monnens LAH, van Oost BA, et al (1994) Mutations in the type IV collagen $\alpha 3$ (*COL4A3*) gene in autosomal recessive Alport syndrome. *Hum Mol Genet* 3: 1269-1273
- Li L, Bray PF (1993) Homologous recombination among three intragene *Alu* sequences causes an inversion-deletion resulting in the hereditary bleeding disorder Glanzmann thrombasthenia. *Am J Hum Genet* 53:140-149
- Maddalena A, Sosnoski DM, Berry GT, Nussbaum RL (1988) Mosaicism for an intragenic deletion in a boy with mild ornithine transcarbamylase deficiency. *N Engl J Med* 319: 999-1003
- Mochizuki T, Lemmink HH, Mariyama M, Antignac C, Gubler MC, Pirson Y, Verellen-Dumoulin C, et al (1994) Identification of mutations in the $\alpha 3$ (IV) and $\alpha 4$ (IV) collagen genes in autosomal recessive Alport syndrome. *Nat Genet* 8:77-82
- Naito I, Kawai S, Nomura S, Sado Y, Osawa G, Matui A, Yoshida M (1996) Relationship between *COL4A5* gene mutation and distribution of type IV collagen in male X-linked Alport syndrome. *Kidney Int* 50:304-311
- Ninomiya Y, Kagawa M, Iyama K, Naito I, Kishiro Y, Seyer JM, Sugimoto M, et al (1995) Differential expression of two basement membrane collagen genes, *COL4A6* and *COL4A5*, demonstrated by immunofluorescence staining using peptide-specific monoclonal antibodies. *J Cell Biol* 130:1219-1229
- Oohashi T, Sugimoto M, Mattei MG, Ninomiya Y (1994) Identification of a new collagen IV chain, $\alpha 6$ (IV), by cDNA isolation and assignment of the gene to chromosome Xq22, which is the same locus for *COL4A5*. *J Biol Chem* 269: 7520-7526
- Oohashi T, Ueki Y, Sugimoto M, Ninomiya Y (1995) Isolation and structure of the *COL4A6* gene encoding the human $\alpha 6$ (IV) collagen chain and comparison with other type IV collagen genes. *J Biol Chem* 270:26863-26867
- Ottolenghi S, Giglioli B (1982) The deletion in a type of δ^0 - β^0 -thalassaemia begins in an inverted *AluI* repeat. *Nature* 300:770-771
- Renieri A, Bruttini M, Galli L, Zanelli P, Neri T, Rossetti S, Turco A, et al (1996) X-linked Alport syndrome: an SSCP-

- based mutation survey over all 51 exons of the COL4A5 gene. *Am J Hum Genet* 58:1192–1204
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2d ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- Schoenmakers EFPM, Wanschura S, Mols R, Bullerdiek J, van den Berghe H, van de Ven WJM (1995) Recurrent rearrangements in the high mobility group protein gene, *HMGC*, in benign mesenchymal tumors. *Nat Genet* 10:436–444
- Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, Cooley T, et al (1987) Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* 1:113–125
- Shapiro LJ, Yen P, Pomerantz D, Martin E, Rolewic L, Mohandas T (1989) Molecular studies of deletions at the human steroid sulfatase locus. *Proc Natl Acad Sci USA* 86: 8477–8481
- Shin RW, Iwaki T, Kitamoto T, Tateishi J (1991) Hydrated autoclave pretreatment enhances TAU immunoreactivity in formalin-fixed normal and Alzheimer's disease brain tissues. *Lab Invest* 64:693–702
- Spitzner JR, Muller MT (1988) A consensus sequence for cleavage by vertebrate DNA topoisomerase II. *Nucleic Acids Res* 16:5533–5556
- Sugimoto M, Oohashi T, Ninomiya Y (1994) The genes *COL4A5* and *COL4A6*, coding for basement membrane collagen chains $\alpha 5$ (IV) and $\alpha 6$ (IV), are located head-to-head in close proximity on human chromosome Xq22 and *COL4A6* is transcribed from two alternative promoters. *Proc Natl Acad Sci USA* 91:11679–11683
- van Bakel I, Holt S, Craig I, Boyd Y (1995) Sequence analysis of the breakpoint region of an X;5 translocation in a female with Duchenne muscular dystrophy. *Am J Hum Genet* 57: 329–336
- Zhou J, Mochizuki T, Smeets H, Antignac C, Laurila P, de Paepe A, Tryggvason K, et al (1993) Deletion of the paired $\alpha 5$ (IV) and $\alpha 6$ (IV) collagen genes in inherited smooth muscle tumors. *Science* 261:1167–1169