Reprint Series
27 August 1993, Volume 261, pp. 1167–1169

SCIENCE

Deletion of the Paired α5(IV) and α6(IV) Collagen Genes in Inherited Smooth Muscle Tumors

Jing Zhou,* Toshio Mochizuki, Hubert Smeets, Corinne Antignac, Pekka Laurila, Anne de Paepe, karl Tryggvason, and Stephen T. Reeders

Deletion of the Paired α5(IV) and α6(IV) Collagen Genes in Inherited Smooth Muscle Tumors

Jing Zhou,* Toshio Mochizuki, Hubert Smeets, Corinne Antignac, Pekka Laurila, Anne de Paepe, Karl Tryggvason, Stephen T. Reeders

The gene encoding α6(IV) collagen, *COL4A6*, was identified on the human X chromosome in a head-to-head arrangement and within 452 base pairs of the α5(IV) collagen gene, *COL4A5*. In earlier studies, intragenic deletions of *COL4A5* were detected in a subset of patients with Alport syndrome (AS), a hereditary defect of basement membranes. In some families, AS cosegregates with diffuse leiomyomatosis (DL), a benign smooth muscle tumor diathesis. Here it is shown that patients with AS-DL harbor deletions that disrupt both *COL4A5* and *COL4A6*. Thus, type IV collagen may regulate smooth muscle differentiation and morphogenesis.

Basement membranes (BMs) compartmentalize tissues and provide important signals for the differentiation of the cells they support. Type IV collagen, the major structural component of BM, is a triple-helical molecule composed of three α chains (*1*). To date, five genetically distinct type IV isoforms have been described in mammals (*1–4*). The α1(IV) and α2(IV) chains encoded by *COL4A1* and *COL4A2*, respectively, are ubiquitous, whereas α3(IV), α4(IV), and α5(IV) have restricted tissue distributions (*3, 5*). On the basis of sequence similarities, the chains fall into two classes: α1(IV), α3(IV), and α5(IV) compose the α1-like class, and α2(IV) and α4(IV) compose the α2-like class. The human genes for the α1(IV) and α2(IV) chains are located in a head-to-head configuration on chromosome 13 (*6*), and the α3(IV) and α4(IV) genes are similarly arranged on chromosome 2 (*7*). Thus, it appears that the type IV collagens evolved by duplication of an ancestral α chain gene, giving rise to a pair of α chain genes with closely apposed 5′ ends. The pair presumably underwent additional duplication, giving rise to the ancestors of the α1(IV)-α2(IV) and α3(IV)-α4(IV) gene pairs. We predicted that the α5(IV) gene, a member of the α1(IV)-like class, might also be paired with an α2(IV)-like gene that had not yet been identified.

Mutations in *COL4A5* are estimated to be present in ~50% of X-linked AS cases (*8*). In males AS is characterized by progressive renal failure, sensorineural deafness, and ocular lesions; female carriers are mildly affected. In some families AS cosegregates with DL (*9*), a benign proliferation of smooth muscle in the esophagus, female genitalia, and trachea. Both sporadic and hereditary cases have been reported (*10*). We have shown that AS-DL patients have deletions that include the 5′ end of *COL4A5* (*11*), whereas AS patients without DL have internal deletions or point mutations of *COL4A5* (*8*). These results suggested that DL is caused by the deletion of an unidentified gene located upstream of *COL4A5*.

To isolate the putative type IV collagen gene upstream of *COL4A5*, we screened an X-chromosome library with JZ-4, an α5(IV) cDNA clone (*4*), and isolated a 14.1-kb clone, λLA226 (Fig. 1). It contained exon 1 of *COL4A5* and an upstream 2.8-kb Hind III fragment, LA226-H6, that displayed cross-species hybridization. We therefore used LA226-H6 to probe an adult kidney cDNA library. Three identical clones, JZK-1, JZK-2, and JZ-3, contained an open reading frame (1643 bp) encoding a 21–amino acid signal peptide, a 25–amino acid noncollagenous segment, and a 502–amino acid collagenous domain with nine interruptions (Fig. 2) that are believed to confer flexibility in type IV collagens. The deduced translation product, which we have termed α6(IV), is a type IV collagen that has not previously been detected genetically or biochemically (*12*). Sequence analysis clearly places α6(IV) in the α2(IV)-like class (Fig. 2). The head-to-head arrangement of *COL4A5* and *COL4A6* resembles that of *COL4A1* and *COL4A2*.

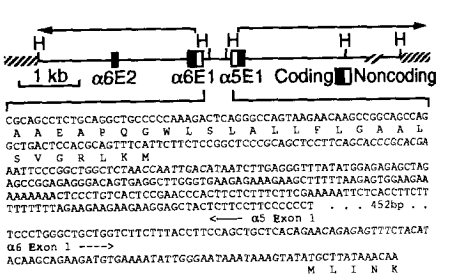

Fig. 1. Restriction map of λLA226 which contains the 5′ ends of both *COL4A6* and *COL4A5*. The striped bars represent phage arms; H, Hind III. The DNA sequence of the region containing the first exon of each gene is shown as well as the deduced amino acid sequences of the first exons (*24*).

J. Zhou, T. Mochizuki, S. T. Reeders, Howard Hughes Medical Institute, Yale University School of Medicine, 295 Congress Avenue, New Haven, CT 06536-0812.
H. Smeets, Clinical Genetics Center, University Hospital, Nijmegen, Netherlands.
C. Antignac, INSERM Unité 192, Hôpital Necker–Enfants Malades, Paris, France.
P. Laurila, Department of Pathology, University of Helsinki, Helsinki, Finland.
A. de Paepe, Department of Human Genetics, University Hospital, Ghent, Belgium.
K. Tryggvason, Biocenter and Department of Biochemistry, University of Oulu, Oulu, Finland.

*To whom correspondence should be addressed.

The relatively low frequency of *COL4A5* mutations in AS patients suggests that another gene may be involved in the X-linked disease. In the adult kidney *COL4A6* is expressed (Fig. 3A) and is therefore a good candidate for a second X-linked AS gene. To explore this possibility, we performed Southern (DNA) blot analysis (Fig. 4) on 140 AS patients, including four unrelated AS-DL patients (*13*), with JZ-3. We saw an abnormal pattern only in the AS-DL patients. There was a loss of bands in males and a 50% reduction in the intensity of some bands in females, but the pattern was complex. We therefore used a fragment, JZ-3-FR5, containing exons 1 to 4 of *COL4A6* to map the deletions more precisely (Fig. 4). The 1.4-kb and 5.6-kb Eco RI fragments, which contain exons 1 and 2 of *COL4A6*, were absent in males and reduced in intensity by ~50% in females; exons 3 and 4 were intact. Hybridization with JZ-4 demonstrated the loss of exon 1 of *COL4A5*. Exons 2 to 10 of *COL4A5* were present in four of the five patients. Therefore, the smallest AS-DL deletions involve part of intron 1 and all of exon 1 of *COL4A5*, the intergenic region, and exons 1 and 2 and part of intron 2 of *COL4A6*.

We studied the distribution of the α5(IV) and α6(IV) transcripts in the fetus (Fig. 3B). An α6(IV) mRNA of ~7.0 kb was found in meninges and esophagus and was just detectable in fetal choroid plexus and stomach. An ~7.0-kb transcript of α5(IV) was present in many tissues but was most abundant in choroid plexus, meninges, and esophagus.

We found deletions of both *COL4A5* and *COL4A6* in all four independent AS-DL kindreds. How do these null mutations cause DL? Interactions between cells and the underlying BM substrate play a vital role in embryonic morphogenesis (*14*). Signals from BM proteins, transduced by members of the integrin family of cell surface receptors, regulate cell growth and differentiation and influence cell shape by affecting the cytoskeleton. Several BM components interact with integrins (*15*) and can affect differentiation of epithelial, endothelial, and mesenchymal cells. Type IV collagens contain binding sites within the triple-helical (*16*) and NC1 domains (*17*) for several cell types including myocytes (*18*). In the presence of antibodies to a β integrin, chicken embryo myoblasts continue to replicate, fail to fuse, and have abnormal morphology (*19*). Although the cited studies are based on collagen sources rich in the ubiquitous α1(IV) and α2(IV) chains (*20*), the data presented here suggest that the α5(IV) and α6(IV) chains may play similar roles in cell-matrix interactions in tissues involved in the AS-DL syndrome; their absence may disrupt normal morphogenesis and lead to uncontrolled proliferation of distorted smooth muscle cells.

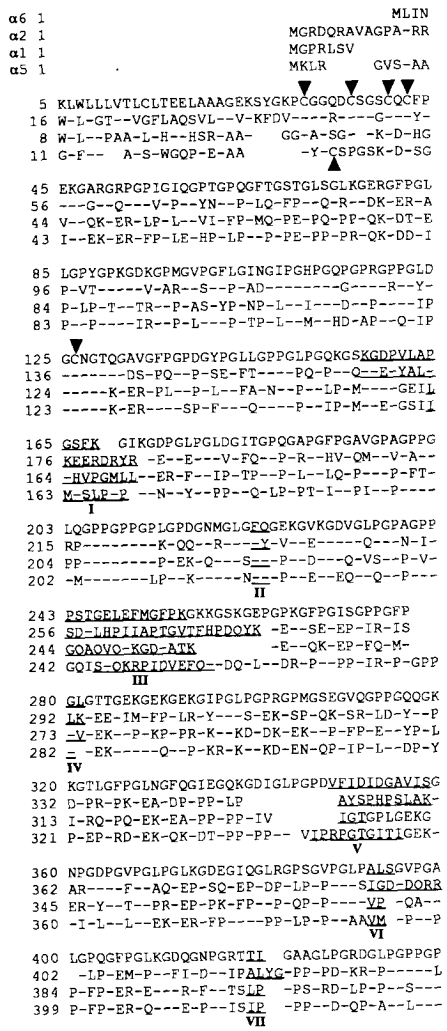

Fig. 2. Deduced amino acid sequences of the cDNA clone JZ-3 of the human α6(IV) chain and comparison of JZ-3 with the amino acid sequences of the human α1(IV), α2(IV), and α5(IV) chains (*4*, *23*). Dashes indicate amino acid identity, and conserved cysteine residues are indicated by arrows. The locations of cysteines in the mature chains are identical in α6(IV) and α2(IV). Interruptions in the collagenous repeat are underlined and numbered (I through IX). The GenBank accession number is L22763.

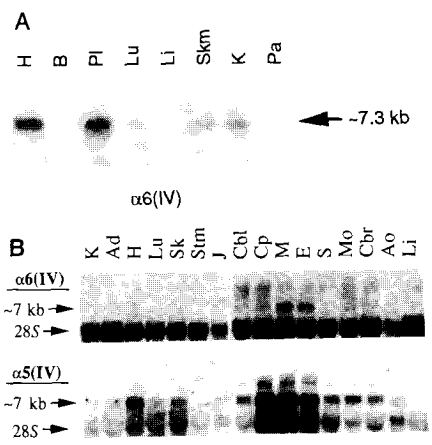

Fig. 3. (**A**) Northern (RNA) blot analysis of tissues from a normal adult. Probe JZ-3, representing α6(IV) sequences, was hybridized to a filter containing ~2 μg of polyadenylate-enriched RNA (per lane) that had been separated by agarose electrophoresis. (**B**) Northern blot analysis of tissues from a 24-week human fetus. Probes JZ-3 and PL-31 were serially hybridized to a filter containing ~20 μg of total RNA in each lane except for samples from aorta (~9 μg) and jejunum (~10 μg). We separated RNAs as in (A). Ad, adrenal; Ao, aorta; B, whole brain; Cbr, cerebrum; Cbl, cerebellum; Cp, choroid plexus; E, esophagus; H, heart; J, jejunum; K, kidney; Li, liver; Lu, lung; M, meninges; Mo, medulla oblongata; Pa, pancreas; PI, term placenta; S, stomach; Skm, skeletal muscle; and Stm, striated muscle.

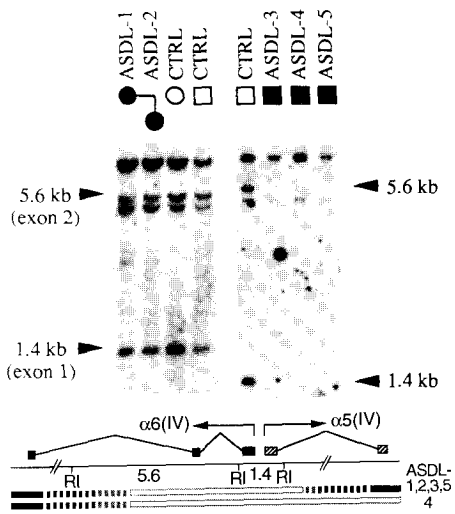

Fig. 4. Analysis of collagen gene deletions in five patients with AS-DL. Genomic DNA samples (~5 μg) from patients (filled symbols) and controls (open symbols) were digested with Eco RI and hybridized with JZ-3-FR5, a cDNA fragment containing the first four exons of α6(IV). Squares indicate males and circles indicate females. In the panel at the bottom, the three exons of α6(IV) and two exons of α5(IV), shown as black and striped boxes, respectively, are placed on a genomic map of Eco RI sites (RI). The open bars below show the minimum extent of the deletion in each AS-DL patient. The filled bars show the minimum extent of the nondeleted regions. The shaded bars represent ambiguities in mapping the deletion boundaries.

Mutations of neither *COL4A5* alone nor *COL4A6* alone have been observed in AS-DL. It is possible that only the α6(IV) chain is critical for normal smooth muscle differentiation. If so, our failure to observe mutations of *COL4A6* alone in AS-DL might merely reflect the small sample size. Linkage studies (*21*) have shown that X-linked AS mutations are all tightly linked to markers in the Xq22 region where both *COL4A5* and *COL4A6* are located. *COL4A6* is the probable site for the ~50% of X-linked AS mutations that have not been found in *COL4A5*. Therefore, the absence of DL in these patients suggests that simultaneous mutation of both *COL4A5* and *COL4A6* is required for the development of DL.

Attention has previously been focused on the role of type IV collagen in adhesion and motility of cells during tumor invasion (*22*). Here we show that constitutional mutations in type IV collagen can result in cell proliferation in benign tumors. Mutations in other BM components, whether constitutional or acquired, may also participate in the pathogenesis of other benign tumors.

REFERENCES AND NOTES

1. R. Timpl, *Eur. J. Biochem.* **180**, 487 (1989).
2. R. J. Butkowski, J. P. M. Langeveld, J. Wieslander, J. Hamilton, B. G. Hudson, *J. Biol. Chem.* **262**, 7874 (1987); K. E. Morrison, M. Mariyama, S. T. Reeders, *Am. J. Hum. Genet.* **49**, 545 (1991); S. Gunwar, J. Saus, M. E. Noelken, B. G. Hudson, *J. Biol. Chem.* **265**, 5466 (1990); M. Mariyama, R. Kalluri, B. G. Hudson, S. T. Reeders, *ibid.* **267**, 1253 (1992); Y. Kamagata, M.-G. Mattei, Y. Ninomiya, *ibid.*, p. 23753; T. Pihlajaniemi, E.-R. Pohjolainen, J. C. Myers, *ibid.* **265**, 13758 (1990).
3. S. L. Hostikka *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1606 (1990).
4. J. Zhou, J. M. Hertz, A. Leinonen, K. Tryggvason, *J. Biol. Chem.* **267**, 12475 (1992).
5. R. J. Butkowski, J. Wieslander, M. Kleppel, A. F. Michael, A. J. Fish, *Kidney Int.* **35**, 1195 (1989); G.-Q. Shen *et al.*, *Connect. Tissue Res.* **24**, 289 (1990); M. Desjardins, F. Gros, J. Wieslander, M. C. Gubler, M. Bendayan, *Lab. Invest.* **63**, 637 (1990); J. R. Sanes, E. Engvall, R. Butkowski, D. D. Hunter, *J. Cell Biol.* **111**, 1685 (1990).
6. R. Soininen, M. Huotari, S. L. Hostikka, D. J. Prockop, K. Tryggvason, *J. Biol. Chem.* **263**, 17217 (1988); E. Pöschl, R. Pollner, K. Kühn, *EMBO J.* **7**, 2687 (1988); P. D. Burbelo, G. R. Martin, Y. Yamada, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9679 (1988).
7. M. Mariyama, K. Zheng, T. L. Yang-Feng, S. T. Reeders, *Genomics* **13**, 809 (1992); M. Mariyama and S. T. Reeders, unpublished data.
8. D. F. Barker *et al.*, *Science* **248**, 1224 (1990); J. Zhou *et al.*, *Genomics* **9**, 10 (1991); K. Tryggvason, J. Zhou, S. L. Hostikka, T. Shows, *Kidney Int.* **43**, 38 (1993); K.-O. Netzer *et al.*, *ibid.* **42**, 1336 (1992).
9. P. Cochat *et al.*, *J. Pediatr.* **113**, 339 (1988).
10. J. B. Marshall, A. A. Diaz-Arias, G. S. Bochna, K. A. Vogele, *Am. J. Gastroenterol.* **85**, 303 (1990).
11. C. Antignac *et al.*, *Kidney Int.* **42**, 1178 (1992).
12. S. Gunwar *et al.*, *J. Biol. Chem.* **266**, 15318 (1991).
13. The patients ASDL-1 and ASDL-2 are related females. ASDL-1 had persistent hematuria and esophageal and genital leiomyomas. Her brother had AS with cataracts and hearing loss and died at age 23 after an unsuccessful renal transplantation. He suffered from an undiagnosed swallowing difficulty and "narrow esophagus." Her mother also had unexplained swallowing difficulty. One sister died at age 37 from myosarcoma of the brain. Two other sisters are healthy. Her daughter, ASDL-2, developed hematuria at age 9. Ultrastructural changes in glomerular basement membrane are typical of AS. She has bilateral cataracts and DL. Patients ASDL-3, -4, and -5 have been described previously (*11*). ASDL-3 is a male with biopsy-proven AS, deafness, cataracts from age 2, and symptomatic esophageal leiomyomatosis since age 9. His mother had persistent hematuria and esophageal and genital leiomyomas. ASDL-4 is a male with AS who developed cataracts and symptomatic esophageal leiomyomatosis at age 4. His mother had hematuria and esophageal leiomyomatosis. ASDL-5 had AS, hematuria, and a grossly dilated esophagus.
14. R. L. Juliano and S. Haskill, *J. Cell Biol.* **120**, 577 (1993).
15. R. O. Hynes, *Cell* **69**, 11 (1992).
16. M. Aumailley and R. Timpl, *J. Cell Biol.* **103**, 1569 (1986).
17. M. K. Chelberg, E. C. Tsilibary, A. R. Hauser, J. B. McCarthy, *Cancer Res.* **49**, 4796 (1989).
18. T. K. Borg, K. Rubin, E. Lundgren, K. Borg, B. Öbrink, *Dev. Biol.* **104**, 86 (1984).
19. A. S. Menko and D. Boettiger, *Cell* **51**, 51 (1987).
20. P. Vandenberg *et al.*, *J. Cell Biol.* **13**, 1475 (1991).
21. H. Brunner *et al.*, *Kidney Int.* **34**, 507 (1988); F. A. Flinter, J. S. Cameron, C. Chantler, I. Houston, M. Bobrow, *Lancet* ii, 1005 (1988).
22. S. Aznavoorian, A. N. Murphy, W. G. Stetler-Stevenson, L. A. Liotta, *Cancer* **71**, 1368 (1993).
23. R. Soininen, T. Haka-Risku, D. J. Prockop, K. Tryggvason, *FEBS Lett.* **225**, 188 (1987); S. L. Hostikka, M. Kurkinen, K. Tryggvason, *ibid.* **216**, 281 (1987).
24. Abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

25. We thank J. Gruen, L. Ollitervo, M. Mariyama, H. Lemmink, and L. Kluytmans for scientific discussion and technical assistance and P. Cochat and B. Roussel for referring patients. Supported by the Sigrid Juselius Foundation, the Academy of Finland, National Institutes of Health, and the Howard Hughes Medical Institute.

22 April 1993; accepted 7 July 1993