**Short communication**

**Contiguous gene deletion breakpoints in Alport syndrome-diffuse leiomyomatosis**

Kandai Nozu1, Shogo Minamikawa1, Shiro Yamada2, Masafumi Oka1, Motoko Yanagita3, Naoya Morisada1, Shuichiro Fujinaga4, China Nagano5, Yoshimitsu Gotoh5, Eihiko Takahashi6, Takahiro Morishita7, Tomohiko Yamamura1, Takeshi Ninchoji1, Hiroshi Kaito1, Ichiro Morioka1, Koichi Nakanishi8, Igor Vorechovsky9 and Kazumoto Iijima1

1. Department of Pediatrics, Kobe University Graduate School of Medicine, Kobe,
2. Department of Pediatrics, Tokai University Oiso Hospital, Oiso, Kanagawa

Division of Human Genetics, National Institute of Genetics, Mishima

1. Department of Nephrology, Kyoto University, Kyoto
2. Division of Nephrology, Saitama Children’s Medical Center, 2100 Magome, Iwatsuki-ku, Saitama City, Saitama
3. Department of Pediatric Nephrology, Japanese Red Cross Nagoya Daini Hospital, Nagoya
4. Department of Nephrology, Kanagawa Children’s Medical Center, Yokohama
5. Department of Pediatrics, University of Occupational and Environmental Health, Fukuoka
6. Department of Pediatrics, Wakayama Medical University, Wakayama
7. University of Southampton Faculty of Medicine, Southampton, United Kingdom

**Running title:** Recombination in AS-DL

**COI Statement:** The authors have nothing to disclose.

**Support sources:** This study was supported by a grant from the Ministry of Health, Labour and Welfare of Japan for Research on Rare Intractable Diseases in Kidney and Urinary Tract (H24-nanchitou (nan)-ippan-041 to Kazumoto Iijima) in the “Research on Measures for Intractable Diseases” Project; a Grant-in-Aid for Scientific Research (KAKENHI) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (Subject ID: 16K19642 to Tomohiko Yamamura, 15K09691 to Kandai Nozu and 26293203 to Kazumoto Iijima).

**Keywords:** Diffuse leiomyomatosis; transposable element; homologous recombination; LINE 1; deletion

**Word count (main text):** 989

**Word count (abstract):** 167

**Corresponding author:**

Kandai Nozu, M.D., Ph.D., Department of Pediatrics, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo, Kobe, Hyogo 6500017, Japan.

Fax: +81-78-382-6099; Tel.: +81-78-382-6090

E-mail: [nozu@med.kobe-u.ac.jp](mailto:nozu@med.kobe-u.ac.jp)**Abstract**

Alport syndrome-diffuse leiomyomatosis (AS-DL, OMIM: 308940) is a rare variant of X-linked Alport syndrome that shows overgrowth of visceral smooth muscles in the gastrointestinal, respiratory, and female reproductive tracts in addition to renal symptoms. AS-DL results from contiguous gene deletions that encompass the 5′ ends of the *COL4A5* and *COL4A6* genes, but the precise deletion breakpoints between *COL4A5* and *COL4A6* have been determined in only four cases. Here, we studied contiguous gene deletion breakpoints in five cases with AS-DL. Including the four previously analyzed cases, we investigated the genetic background that causes these deletions. We demonstrate that eight out of nine deletion alleles involved sequences homologous between *COL4A5* and *COL4A6*. Most breakpoints took place in recognizable transposed elements, including long interspersed repeats, DNA transposons and long-terminal repeat retrotranposons. Deletions involved the bidirectional promoter region in each case, suggesting that the occurrence of leiomyomatosis in AS-DL requires inactivation of both genes. Together, our study highlights the importance of homologous recombination involving multiple transposed elements for the development of atypical loss-of-function phenotypes. **Introduction**

Alport syndrome (AS) is the most common hereditary nephropathy, characterized by progressive renal failure, sensorineural deafness, and ocular abnormalities. The most common X-linked form of this disease (XLAS, OMIM: 301050) results from mutations in the *COL4A5* gene, which encodes the α5 chain of type IV collagen [1](#_ENREF_1). Alport syndrome-diffuse leiomyomatosis (AS-DL, OMIM: 308940) occurs as a rare variant of XLAS that shows overgrowth of visceral smooth muscles in the gastrointestinal, respiratory, and female reproductive tracts, in addition to renal symptoms [1-3](#_ENREF_1). *COL4A5* is located on the long arm of the X chromosome (Xq22) and is head-to-head with another type IV collagen gene, *COL4A6.* The *COL4A6* gene encodes the α6 chain of type IV collagen that is mainly expressed in heart, human esophagus, aorta, and bladder smooth muscle basement membrane[3-5](#_ENREF_3) together with the α5 chain.

AS-DL patients typically exhibit contiguous gene deletions at the *COL4A5–COL4A6* locus [5](#_ENREF_5). Sixteen patients reported so far with AS-DL have been found to have a deletion that encompasses the 5′ end of the *COL4A5* and *COL4A6* genes, and includes the bidirectional promoter (according to the HGMD data base). The *COL4A6* deletion breakpoints have been consistently found within intron 2, whereas the *COL4A5* breakpoints usually occur in intron 1 [6](#_ENREF_6). However, the precise breakpoints have been identified by direct sequencing in only four cases in previous studies [6-9](#_ENREF_6).

**Patients and Methods**

In our laboratory, we conduct the comprehensive analysis for inherited kidney diseases including Alport syndrome. More than 400 suspected Alport syndrome patients’ samples were applied for the genetic tests. Among them, from the clinical datasheets, five of them were shown to be complicated by leiomyomatosis. For all five of them, MLPA was conducted and detected contiguous gene deletions, and then, included to this study. The pedigrees of the five cases are shown in SupplementaryFigure 1. All cases showed esophageal leiomyomatosis and nephropathy. The clinical information of all cases are shown in Supplementary Table 1.

Genomic DNA was isolated from peripheral blood leukocytes. Screening of contiguous gene deletions was performed using the SALSA P191/192 Alport MLPA assay (MRC-Holland, Amsterdam, the Netherlands) as indicated by the manufacturer[2](#_ENREF_2), [10](#_ENREF_10), [11](#_ENREF_11). Long-range PCR amplification and sequencing of *COL4A5* and *COL4A6* and direct sequencing was conducted to detect the breakpoint of the contiguous gene deletions.

To detect a large heterozygous deletion in case 2, we conducted semi-quantitative PCR amplification using capillary electrophoresis[12](#_ENREF_12), [13](#_ENREF_13). The *COL4A5* and *COL4A6* reads were mapped to the human reference sequence NG\_011977.1 and NG\_012059.2, respectively.

**Results**

The MLPA analysis of five patients with AS-DL revealed hetero- or hemizygous deletions in each case (Supplementary Figure 2). This analysis was followed by **l**ong-range PCR and direct sequencing to identify each breakpoint at the highest resolution (Figure 1A–E, Supplementary Figure 3, and Supplementary Data 1). For a female patient (Case 2, SupplementaryFigure 2B), we also conducted a semi-quantitative PCR assay (SupplementaryFigure 4) before mapping the breakpoints at the single nucleotide level (Figure 1). Deletions are schematically shown in SupplementaryFigure 3, which includes four previously reported cases of AS-DL (p1–p4) [6-9](#_ENREF_6) for comparison. To examine the genomic context of deletion breakpoints, we analyzed their flanking sequences using RepeatMasker (Figure 2). Their alignments with Repbase entries [14](#_ENREF_14) revealed the presence of transposed elements (TEs) in four of our five cases and in 3 of 4 recombination events in previously reported AS-DL patients (Figure 2). Cases p2 and p3 showed *COL4A5*-side breakpoints at different positions within the same long interspersed element L1PA3. Cases c3 and p2 exhibited *COL4A6*-side breakpoints at different positions in the same L1P1 copy that were separated by 562 nucleotides. Reference sequences flanking deletion breakpoints are shown in Supplementary data 1. Homologous sequences at the recombination breakpoints were apparent in eight of nine patients. The length of overlapping sequences varied from 2 to 42 bp (Figure 2). We have conducted MLPA for the mothers of Case 1, 3 and 4 and could not detect the contiguous gene deletions in those samples. Strictly speaking, the father’s sample should be analyzed for the female case (Case 3). However, he doesn’t have urine abnormalities or leiomyomatosis. Therefore, all these three cases are assumed to be sporadic cases (Supplementary Figure 1 and Supplementary Table 1).

**Discussion**

This work has more than doubled the number of sequence-characterized *COL4A5* and *COL4A6* breakpoints in AS-DL. Although eight out of nine cases had homologous sequences at the deletion breakpoints, only two cases (c3 and p2) showed relatively long homologous sequences in L1 retrotransposons, consistent with a well-known L1-mediated recombination mechanism [15](#_ENREF_15). Almost a half of the human genome is occupied by recognizable TEs [16](#_ENREF_16). TEs have been shown to provide a source of new exons or genes, dramatically influencing evolutionary history, exon–intron structure, speciation and regulation of gene expression. TEs facilitate non-allelic homologous recombination events leading to inherited diseases [17](#_ENREF_17), including AS-DL (Figure 4) [8](#_ENREF_8). Our results show that the deletions in c3 and p2 resulted from a homologous recombination between L1 elements, but additional cases (c1, c4, c5, p1, and p3) revealed shorter homologous sequences indicative of the same mechanism. The recombination breakpoints tend to occur in the same TE locus, as illustrated for cases c3 and p2. These regions may represent hotspots for homologous recombination events that lead to AS-DL.

All deletion breakpoints characterized at the single nucleotide level took place in *COL4A6* intron 2 [6-9](#_ENREF_6). Recently, an AS-DL case was identified with a deletion extending beyond this intron [2](#_ENREF_2), but the precise breakpoint was not characterized. On the *COL4A5* side, the deletion breakpoint usually maps to intron 1; however, some cases, including two of ours (c3 and c4, Figure 4), showed more centromeric 3′ breakpoints, even at the 5′ end of *COL4A5*. Recently, Sa et al. reported an AS-DL case with a *COL4A5*-only gene deletion, which encompassed exons 1 to 51 but did not include the common promoter region or exon 1 of *COL4A6* [18](#_ENREF_18). This report may still be compatible with the requirement for inactivation of both genes is required for the AS-DL phenotype, as first proposed by Zhou et al. [5](#_ENREF_5). The authors did not analyze *COL4A6* expression nor did they exclude inactivation of this gene by other mechanisms. Recent studies in yeasts revealed that deletion of many genes were associated with altered mRNA levels of the neighboring genes. These studies are particularly relevant for bidirectional promoters which generate protein products from two adjacent related genes in stoichiometric quantities, ensuring co-expression of genes in the same or similar biological pathways.

In conclusion, we have more than doubled the number of large contiguous deletions in AS-DL characterized at the single nucleotide level (Figure 2). Our results show that most deletions were mediated by transposons via homologous recombination events and support the original proposal [5](#_ENREF_5) that inactivation of both genes is required for the development of leiomyomas in AS-DL.

**COI Statement:** The authors have nothing to disclose.

**Acknowledgments** The authors gratefully acknowledge the cooperation of the patients and physicians. This study was supported by a grant from the Ministry of Health, Labour and Welfare of Japan for Research on Rare Intractable Diseases in Kidney and Urinary Tract (H24-nanchitou (nan)-ippan-041 to Kazumoto Iijima) in the “Research on Measures for Intractable Diseases” Project; a Grant-in-Aid for Scientific Research (KAKENHI) from the Ministry of Education, Culture, Sports, Science and Technology of Japan (Subject ID: 16K19642 to Tomohiko Yamamura, 25893131 to Kandai Nozu and 26293203 to Kazumoto Iijima).

**References**

1. Kashtan, C.E. Alport syndrome and thin glomerular basement membrane disease. *J Am Soc Nephrol.* **9**, 1736-1750 (1998).

2. Uliana, V., Marcocci, E., Mucciolo, M., Meloni, I., Izzi, C., Manno, C. et al. Alport syndrome and leiomyomatosis: the first deletion extending beyond COL4A6 intron 2. *Pediatr Nephrol.* **26**, 717-724 (2011).

3. Van Loo, A., Vanholder, R., Buytaert, I., De Paepe, A., Praet, M., Elewaut, A. et al. Alport syndrome and diffuse leiomyomatosis with major morbid events presenting at adult age. *Nephrol Dial Transplant.* **12**, 776-780 (1997).

4. Heidet, L., Cai, Y., Sado, Y., Ninomiya, Y., Thorner, P., Guicharnaud, L. et al. Diffuse leiomyomatosis associated with X-linked Alport syndrome: extracellular matrix study using immunohistochemistry and in situ hybridization. *Lab Invest.* **76**, 233-243 (1997).

5. Zhou, J., Mochizuki, T., Smeets, H., Antignac, C., Laurila, P., de Paepe, A. et al. Deletion of the paired alpha 5(IV) and alpha 6(IV) collagen genes in inherited smooth muscle tumors. *Science.* **261**, 1167-1169 (1993).

6. Oohashi, T., Naito, I., Ueki, Y., Yamatsuji, T., Permpoon, R., Tanaka, N. et al. Clonal overgrowth of esophageal smooth muscle cells in diffuse leiomyomatosis-Alport syndrome caused by partial deletion in COL4A5 and COL4A6 genes. *Matrix Biol.* **30**, 3-8 (2011).

7. Thielen, B.K., Barker, D.F., Nelson, R.D., Zhou, J., Kren, S.M. & Segal, Y. Deletion mapping in Alport syndrome and Alport syndrome-diffuse leiomyomatosis reveals potential mechanisms of visceral smooth muscle overgrowth. *Hum Mutat.* **22**, 419 (2003).

8. Segal, Y., Peissel, B., Renieri, A., de Marchi, M., Ballabio, A., Pei, Y. et al. LINE-1 elements at the sites of molecular rearrangements in Alport syndrome-diffuse leiomyomatosis. *Am J Hum Genet.* **64**, 62-69 (1999).

9. Ueki, Y., Naito, I., Oohashi, T., Sugimoto, M., Seki, T., Yoshioka, H. et al. 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. *Am J Hum Genet.* **62**, 253-261 (1998).

10. Nozu, K., Krol, R.P., Nakanishi, K., Yoshikawa, N., Nozu, Y., Ohtsuka, Y. et al. Detection by multiplex ligation-dependent probe amplification of large deletion mutations in the COL4A5 gene in female patients with Alport syndrome. *Pediatr Nephrol.* **24**, 1773-1774 (2009).

11. Hertz, J.M., Juncker, I. & Marcussen, N. MLPA and cDNA analysis improves COL4A5 mutation detection in X-linked Alport syndrome. *Clin Genet.* **74**, 522-530 (2008).

12. Nozu, K., Przybyslaw Krol, R., Ohtsuka, Y., Nakanishi, K., Yoshikawa, N., Nozu, Y. et al. Detection of large deletion mutations in the COL4A5 gene of female Alport syndrome patients. *Pediatr Nephrol.* **23**, 2085-2090 (2008).

13. Nozu, K., Fu, X.J., Nakanishi, K., Yoshikawa, N., Kaito, H., Kanda, K. et al. Molecular analysis of patients with type III Bartter syndrome: picking up large heterozygous deletions with semiquantitative PCR. *Pediatr Res.* **62**, 364-369 (2007).

14. Tempel, S., Jurka, M. & Jurka, J. VisualRepbase: an interface for the study of occurrences of transposable element families. *BMC Bioinformatics.* **9**, 345 (2008).

15. Gilbert, N., Lutz-Prigge, S. & Moran, J.V. Genomic deletions created upon LINE-1 retrotransposition. *Cell.* **110**, 315-325 (2002).

16. Lander, E.S. & Linton, L.M. & Birren, B. & Nusbaum, C. & Zody, M.C. & Baldwin, J. et al. Initial sequencing and analysis of the human genome. *Nature.* **409**, 860-921 (2001).

17. Belancio, V.P., Hedges, D.J. & Deininger, P. Mammalian non-LTR retrotransposons: for better or worse, in sickness and in health. *Genome Res.* **18**, 343-358 (2008).

18. Sa, M.J., Fieremans, N., de Brouwer, A.P., Sousa, R., e Costa, F.T., Brito, M.J. et al. Deletion of the 5'exons of COL4A6 is not needed for the development of diffuse leiomyomatosis in patients with Alport syndrome. *J Med Genet.* **50**, 745-753 (2013).

**Titles and Legends to Figures**

**Figure 1.** **Sequencing chromatograms for the contiguous gene deletion breakpoints in all four cases**

A, Case 1 ( c.66 + 5840 of *COL4A6* and c.81 + 8068 of *COL4A5*), B,

Case 2 (c.66 + 25107 of *COL4A6* and c.81 + 18040 of *COL4A5*). C, Case 3 (c.66 + 85676 of *COL4A6* and c.276 + 3257 of *COL4A5*). D, Case 4 (c.66 + 119476 of *COL4A6* and c.3246 + 6706 of *COL4A5*). E, Case 5 (c.66 + 84055 of *COL4A6* and c.3246 + 66915 of *COL4A5*). Here, *COL4A6* and *COL4A5* sequences are shown as black and open rectangles, respectively. Homologous sequence in each case is shown in red.

**Figure 2. Schematics of novel and previously reported deletions Location of both sides of the breakpoints in the in *COL4A5* and *COL4A6***

Deletions are shown as dark rectangle. Transposable elements in cenetromeric (*COL4A6*) and telomeric (*COL4A5*) breakpoints identified by RepeateMasker (version 3.0) are shown in the table next to the figure. The first two exons of each gene are numbered at the top. c1-c5, our case 1-5, p1 (ref 4), p2 (ref 17), p3 (ref 14), p4 (ref 13). The nucleotide numbers of overlaps (nucleotides) are shown in parentheses.