Detecting MUC1 Variants in Patients Clinicopathologically Diagnosed With Having Autosomal Dominant Tubulointerstitial Kidney Disease

Okada, Eri ; Morisada, Naoya ; Horinouchi, Tomoko ; Fujii, Hideki ; Tsuji, Takayuki ; Miura, Masayoshi ; Katori, Hideyuki ; Kitagawa,…

(Citation)
Kidney International Reports,7(4):857-866

(Issue Date)
2022-04

(Resource Type)
journal article

(Version)
Version of Record

(URL)
https://hdl.handle.net/20.500.14094/90009534

KI REPORTS
KIReports.org
CLINICAL RESEARCH

Detecting *MUC1* Variants in Patients Clinicopathologically Diagnosed With Having Autosomal Dominant Tubulointerstitial Kidney Disease

Eri Okada[1,2], Naoya Morisada[1,3], Tomoko Horinouchi[1], Hideki Fujii[4], Takayuki Tsuji[5], Masayoshi Miura[6], Hideyuki Katori[7], Masashi Kitagawa[8], Kunio Morozumi[9], Takanobu Toriyama[10], Yuki Nakamura[11], Ryuta Nishikomori[12], Sadayuki Nagai[1], Atsushi Kondo[1], Yuya Aoto[1], Shinya Ishiko[1], Rini Rossanti[1], Nana Sakakibara[1], China Nagano[1], Tomohiko Yamamura[1], Shingo Ishimori[1], Joichi Usui[2], Kunihiro Yamagata[2], Kazumoto Iijima[13,14], Toshiyuki Imasawa[15] and Kandai Nozu[1]

[1]Department of Pediatrics, Kobe University Graduate School of Medicine, Kobe, Japan; [2]Department of Nephrology, Faculty of Medicine, University of Tsukuba, Tsukuba, Japan; [3]Department of Clinical Genetics, Hyogo Prefectural Kobe Children's Hospital, Kobe, Japan; [4]Division of Nephrology and Kidney Center, Kobe University Graduate School of Medicine, Kobe, Japan; [5]First Department of Medicine, Hamamatsu University School of Medicine, Hamamatsu, Japan; [6]Department of Urology and Renal Transplantation Surgery, Sapporo Hokuyu Hospital, Sapporo, Japan; [7]Department of Nephrology, Ebina General Hospital, Ebina, Japan; [8]Department of Nephrology, Rheumatology, Endocrinology and Metabolism, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan; [9]Department of Nephrology, Masuko Memorial Hospital, Nagoya, Japan; [10]Toyohashi Mates Clinic, Toyohashi, Japan; [11]Department of Nephrology and Rheumatology, Iwate Prefectural Central Hospital, Morioka, Japan; [12]Department of Pediatrics and Child Health, Kurume University School of Medicine, Kurume, Japan; [13]Hyogo Prefectural Kobe Children's Hospital, Kobe, Japan; [14]Department of Advanced Pediatric Medicine, Kobe University Graduate School of Medicine, Kobe, Japan; and [15]Division of Nephrology, National Hospital Organization Chiba-Higashi National Hospital, Chiba, Japan

Introduction: Autosomal dominant tubulointerstitial kidney disease (ADTKD)-*MUC1* is predominantly caused by frameshift mutations owing to a single-base insertion into the variable number tandem repeat (VNTR) region in *MUC1*. Because of the complexity of the variant hotspot, identification using short-read sequencers (SRSs) is challenging. Although recent studies have revealed the usefulness of long-read sequencers (LRSs), the prevalence of *MUC1* variants in patients with clinically suspected ADTKD remains unknown. We aimed to clarify this prevalence and the genetic characteristics and clinical manifestations of ADTKD-*MUC1* in a Japanese population using an SRS and an LRS.

Methods: From January 2015 to December 2019, genetic analysis was performed using an SRS in 48 patients with clinically suspected ADTKD. Additional analyses were conducted using an LRS in patients with negative SRS results.

Results: Short-read sequencing results revealed *MUC1* variants in 1 patient harboring a cytosine insertion in the second repeat unit of the VNTR region; however, deeper VNTR regions could not be read by the SRS. Therefore, we conducted long-read sequencing analysis of 39 cases and detected *MUC1* VNTR variants in 8 patients (in total, 9 patients from unrelated families). With the inclusion of family-affected patients ($n = 31$), the median age at the development of end-stage kidney disease (ESKD) was 45 years (95% CI: 40–40 years).

Conclusion: In Japan, the detection rate of *MUC1* variants in patients with clinically suspected ADTKD was 18.8%. More than 20% of patients with negative SRS results had *MUC1* variants detected by an LRS.

Kidney Int Rep (2022) **7,** 857–866; https://doi.org/10.1016/j.ekir.2021.12.037
KEYWORDS: ADTKD; ADTKD-*MUC1*; long-read sequencing; MCKD; NGS; SMRT sequencing

Correspondence: Eri Okada, Department of Pediatrics, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan. E-mail: s2030356@s.tsukuba.ac.jp
Received 19 October 2021; revised 15 December 2021; accepted 28 December 2021; published online 4 January 2022

ADTKD is a group of hereditary kidney diseases characterized by progressive tubulointerstitial fibrosis and tubular atrophy leading to ESKD.[1] ADTKD can be divided into subtypes based on the underlying causative genetic defects,

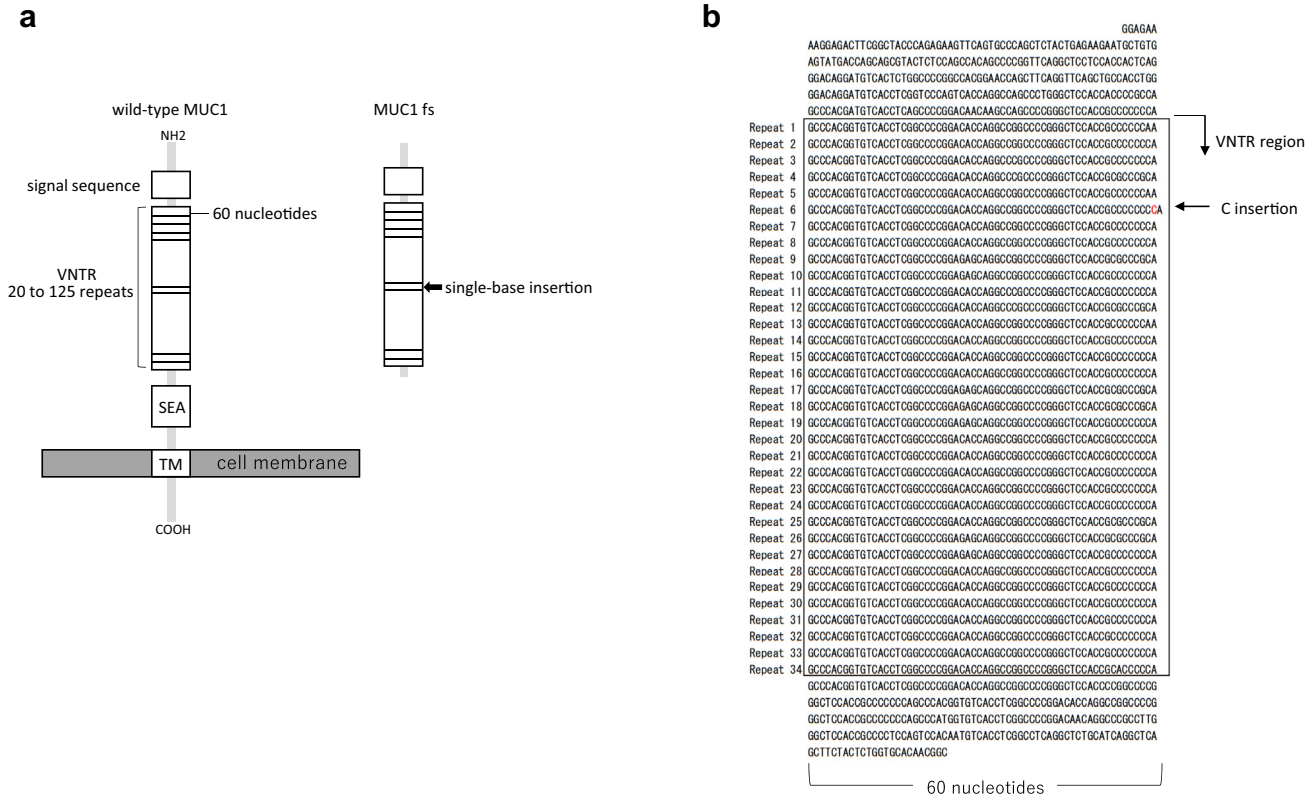

Figure 1. (a) Schemas of wild-type *MUC1* (left) and mutated *MUC1* (MUC1fs, right). Wild-type *MUC1* contains the signal sequence, VNTR region, SEA domain, and TM domain. The VNTR region is composed of 60-nucleotide units repeated 20 to 125 times (each allele has unique repeats). A single-base insertion into the VNTR region terminates synthesis of MUC1 protein after the VNTR region and creates MUC1fs, which lacks a C-terminal domain. (b) Representative sequence of the VNTR region harboring a cytosine insertion in the sixth repeat (SC416). SEA, sea urchin sperm protein-enterokinase-agrin; TM, transmembrane; VNTR, variable tandem repeat.

including abnormalities in *UMOD*, *MUC1*, *REN*, *HNF1B*, and *SEC61A1*.[1,2] Given the nonspecificity of the common unifying characteristics of ADTKD, including bland urinary sediment abnormalities and mild to negative proteinuria, ADTKD is presumed to be highly underdiagnosed.[3]

MUC1 is a transmembrane glycoprotein highly expressed in the apical membrane of the thick ascending limb of the loop of Henle, distal convoluted tubules, and collecting ducts.[4] The most common genetic cause of ADTKD-*MUC1* is a frameshift variant caused by the insertion of a single base into the VNTR polymorphism of *MUC1* (OMIM 158340; 1q22) in exon 2.[5] This mutation terminates the synthesis of the MUC1 protein after the VNTR region and creates a new shortened protein (MUC1fs) that lacks a C-terminal domain (Figure 1a). MUC1fs then accumulates in the cytoplasm owing to abnormal intracellular protein transport. The VNTR region exhibits substantial interindividual variation because it is composed of 60-nucleotide units that repeat 20 to 125 times, with each allele having unique repeats (Figure 1b).[5] Therefore, detecting gene variants using SRSs has been challenging owing to the complexity of the variant hotspot.

Previous studies have used alternative methods, including mass spectrometry[5] and probe extension assays,[6] to detect causative variants; however, these methods are technically demanding. Recent studies have found the usefulness of LRSs, which use a single-molecule real-time (SMRT) sequencing method to identify *MUC1* VNTR variants.[7,8] Nevertheless, *MUC1* VNTR variants have not yet been characterized in the Japanese population. In addition, the prevalence of *MUC1* variants in patients with clinically suspected ADTKD remains unknown. Therefore, the aim of this study was to clarify this prevalence, including the genetic characteristics and clinical manifestations, of ADTKD-*MUC1* in the Japanese population. Here, we diagnosed 9 patients with ADTKD-*MUC1* who had VNTR variants detected using either an SRS ($n = 1$) or an LRS ($n = 8$) among 48 clinically suspected ADTKD cases.

METHODS

Participants

Between January 2015 and December 2019, we performed genetic analysis in 48 patients with clinically suspected ADTKD. The participants had renal

dysfunction from unknown causes, with mild to negative urinary abnormalities. Patients were clinically diagnosed with having ADTKD if they had a positive family history of chronic kidney disease and/or hyperuricemia or if pathologic findings on kidney biopsy revealed either medullary cystic kidney disease (the previous term for ADTKD-*MUC1* and ADTKD-*UMOD*), tubulointerstitial nephritis, or chronic interstitial damage. We excluded patients with renal malformations and/or extrarenal symptoms. DNA samples were first analyzed using an SRS. Further investigations were then performed in patients with negative SRS results ($n = 39$) using an LRS. All samples and clinical information were obtained from Japanese referral hospitals. Genetic analyses were performed after obtaining written informed consent from patients and/or their guardians if the patient was a minor. This study was approved by the Institutional Review Board of the Kobe University School of Medicine (approval number 301).

Short-Read Sequencing

Genomic DNA was extracted from peripheral blood leukocytes using the QuickGene Mini 80 system (Wako Pure Chemical Industries, Ltd., Tokyo, Japan). For targeted sequencing using an SRS, samples were prepared using HaloPlex (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. Paired-end sequencing was performed on the MiSeq platform (Illumina, San Diego, CA). HaloPlex was used for targeted sequencing of 128 genes (version 2, Supplementary Table S1), 172 genes (version 4, Supplementary Table S2), 159 genes (version 5, Supplementary Table S3), 164 genes (version 6, Supplementary Table S4), 181 genes (version 7, Supplementary Table S5), and 183 genes (version 8, Supplementary Table S6) associated with congenital anomalies of the kidney and urinary tract, cystic kidneys, ADTKD, and nephronophthisis. The HaloPlex version used for each patient is found in Supplementary Table S7. Reads were aligned to the reference human genome (GRCh37/Hg19) using SureCall 4.0, which is a desktop application combining algorithms for end-to-end next-generation sequencing data analysis from alignment to categorization of mutations (Agilent Technologies). We excluded called variants with minor allele frequencies >1% in publicly available human variation databases, such as the following: Human Genetic Variation Database (http://www.genome.med.kyoto-u.ac.jp/SnpDB/), gnomAD (https://gnomad.broadinstitute.org/), and 1000 Genomes Project (1000G) databases (https://www.internationalgenome.org/data). Candidate variants were selected using the computational prediction software, SIFT (https://sift.bii.astar.edu.sg/), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), Mutation Taster (http://www.mutationtaster.org/), and CADD (https://cadd.gs.washington.edu/snv), and then classified as pathogenic, likely pathogenic, or uncertain significance, according to the guidelines of the American College of Medical Genetics and Genomics.[9]

Sanger Sequencing

MUC1 variants detected by SRS were confirmed using Sanger direct sequencing. For sequencing, polymerase chain reactions (PCRs) were performed using KOD-Multi & Epi (Toyobo Inc., Osaka, Japan) as the DNA polymerase with the following primer pairs: forward, 5′-GCTGCTCCTCACAGTGCTTA-3′; reverse, 5′-AGGAGGTACCGTGCTATGGT-3′. The PCR conditions were as follows: initial denaturation at 94 °C for 2 minutes followed by 30 PCR cycles at 98 °C for 10 seconds, 58 °C for 10 seconds, and 68 °C for 40 seconds. PCR products were sequenced using 5′-GAAGTTCAGTGCCCAGCTCT-3′ with the 3130 Genetic Analyzer (Thermo Fisher Scientific, Waltham, MA).

PCR Amplification of the *MUC1* VNTR Region for Long-Read Sequencing

Primers were designed for each sample according to a previous study[7] and the guidelines for using PacBio Barcodes for SMRT Sequencing (Pacific Biosciences, Menlo Park, CA), with 5′-XXXXXXXXXXXXXXXGGAGAAAAGGAGACTTCGGCTACCCAG-3′ as the forward primer and 5′-XXXXXXXXXXXXXXXGCCGTTGTGCACCAGAGTAGAAGCTGA-3′ as the reverse primer (X represents each nucleotide in PacBio Barcodes V2). The PCR reaction mixture contained 50 ng of genomic DNA, 5 pmol of primers, 0.5 U of KOD Multi & Epi (Toyobo Inc., Osaka, Japan), 10 µl of 2× PCR buffer for KOD Multi & Epi, 8 ml of double-distilled water, and 2 µl of dimethylsulfoxide for a total volume of 22 µl. The PCR amplification reaction was performed by initial denaturation (2 minutes at 94 °C) and 25 cycles of amplification (10 seconds at 98 °C, 10 seconds at 63 °C, and 4 minutes and 30 seconds at 68 °C) using a Veriti 96-well thermal cycler (Thermo Fisher Scientific, Waltham, MA). The PCR products were analyzed by electrophoresis using an Agilent 2100 Bioanalyzer (Agilent Technologies) with a high-sensitivity DNA kit. Amplicons were extracted and purified using a MinElute PCR purification kit (QIAGEN, Venlo, The Netherlands).

Library Preparation and Long-Read Sequencing (SMRT Sequencing)

The amplicons were pooled in equimolar ratios and prepared for SMRTbell sequence libraries using a SMRTbell Express Template Prep Kit 2.0 (Pacific

Biosciences, Menlo Park, CA) according to the manufacturer's instructions. Subsequently, a sequence template was constructed by attaching the sequence primer version 2 and DNA polymerase to the adapters at both ends of the sequence library, and the template was loaded onto an SMRT cell. SMRT sequencing was performed using the PacBio Sequel II System with Sequel II Sequencing Kit 2.0. The sequenced data were analyzed with SMRT Link version 9.0.0, and highly accurate consensus sequences were obtained. Thereafter, sequence clustering was conducted using the analysis tool "pbaa," which is suitable for analyzing repetitive sequences. Library preparation, sequencing, and bioinformatics analyses of SMRT sequencing were conducted by Takara Bio (Kusatsu, Shiga, Japan). The nucleotide sequence, number of reads, and allele frequency were compiled into an Excel file for each patient (data not revealed). For each identified sequence, we manually reconstructed the sequence every 60 base per line, as described by Kirby *et al.*[5] (Figure 1b). Because of potential PCR errors in the complex VNTR sequence, multiple sequences were detected in each patient. The correct sequences were identified by comparing the LRS results with electropherograms of the amplicons obtained using a bioanalyzer (Supplementary Table S8). In most cases, the top 2 sequences with the highest number of reads were determined to be correct. In some cases, however, sequences with the third or lower most common number of reads were determined to be correct if their sizes were matched to the electrophoresis peaks detected by the bioanalyzer (SC449, SC511, SC534, SC560, SC566, SC593, SC616, and SC639) owing to the inefficiency of PCR amplification of long alleles when the size of the VNTR region of the 2 alleles was significantly different.

Statistical Analysis

All statistical analyses were performed with EZR,[10] which is a modified version of R commander designed to include the statistical functions frequently used in biostatistics. Kaplan–Meier curves were generated for the time taken to develop ESKD.

RESULTS

Identification of Pathogenic Variants by SRS and LRS

Variants in the following genes were detected during initial SRS screening: *UMOD* ($n = 6$), *REN* ($n = 1$), *NPHP1* ($n = 1$), and *MUC1* ($n = 1$) (variant details and clinicopathologic characteristics of patients with SRS-detected variants, other than MUC1, are presented in Supplementary Table S9 and S10, respectively). We conducted further investigations using LRS for 39 patients with negative results on initial SRS screening.

Table 1. Number of study participants and detected gene variants

	n = 48		
Gene	SRS ($n = 48$)	LRS for *MUC1* ($n = 39$)	Total
UMOD	6	–	6
MUC1	1	8	9
REN	1	–	1
NPHP1	1	–	1

LRS, long-read sequencer; SRS, short-read sequencer.

We identified *MUC1* VNTR variants in 8 of these patients, resulting in the identification of *MUC1* variants in a total of 9 patients from unrelated families. Therefore, the detection rate of *MUC1* variants using SRS and LRS in patients with clinically suspected ADTKD was 18.8% (9 of 48 patients) (Table 1).

Variant Details of Detected *MUC1* Mutations

Using SRS with Integrative Genomics Viewer version 2.5.0 (Figure 2a)[11] and Sanger sequencing (NM_001204285.1: c.401_402insC, p.Ala135Serfs*86) (Figure 2b), we identified a single cytosine insertion into 7 consecutive cytosines at the end of the 60-nucleotide unit in the second repeat of the VNTR sequence presented in Figure 1b. An SRS was used to annotate the alignment of VNTR until the third repeat of the whole repeats. In patients with LRS-detected variants, repeat numbers of the 60-nucleotide unit in VNTR harboring an inserted nucleotide ranged from 5 to 40. There were 9 patients (SC359, SC696, SC798, SC416, SC489, SC534, SC656, SC512, and SC566) who exhibited the same 60-nucleotide unit sequence harboring a cytosine. This mutation was reported as "27dupC" by Živná *et al.*[12] using a MASS spectrometry-based probe extension assay. A novel mutation was identified in the 2 remaining patients (SC356 and SC370). The variant detected in SC356 was also a cytosine insertion at the end of the 60-nucleotide unit; however, the original sequence differed in the 29th nucleotide (C to A), which resulted in a non-synonymous variant (threonine to serine). Patient SC370 had a guanine insertion in the middle part of the 60-nucleotide unit (Table 2).

Identification of Alleles and Confirmation of the Repeat Number of the 60-Nucleotide Unit in VNTR

Among 39 patients who underwent genetic testing with an LRS, 36 patients had heterozygous repeat numbers of VNTR and 2 patients (SC515 and SC634) had homozygous repeat numbers (Supplementary Table S11). The VNTR repeat number of each allele ranged from 30 to 82, with 79% containing 30 to 36 repeats (Supplementary Figure S1).

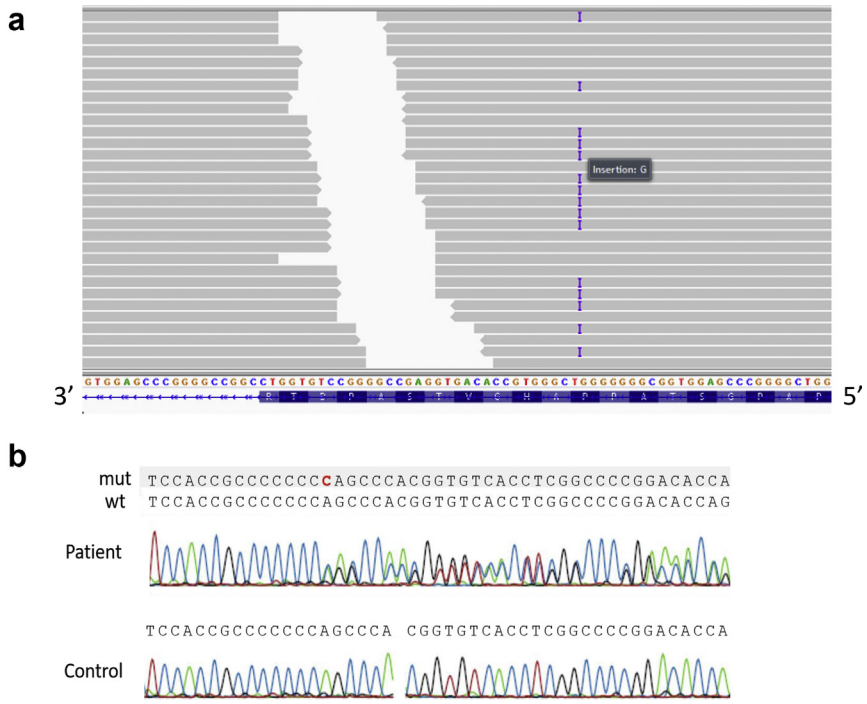

Figure 2. Genetic testing using next-generation sequencing revealed a cytosine insertion in the VNTR region of *MUC1*. (a) Single cytosine insertion in the VNTR region was identified using Integrative Genomics Viewer version 2.7.2 (Broad Institute). Because the reference sequence and each read are displayed as complementary strands, the right side of the figure represents the 5′-end and the left side represents the 3′-end. Inserted nucleotide was displayed as complementary "G" instead of "C." (b) The insertion was confirmed by Sanger sequencing. VNTR, variable tandem repeat.

Clinical Manifestation of Patients Harboring *MUC1* Mutations

Clinical and histopathologic data and variant details are summarized in Table 3. The median age of the 9 patients was 42 years (interquartile range: 39–40 years). There were 6 patients who were of male sex and 3 of female sex. Most of the patients had a family history of ESKD, except 2 (SC512 and SC566). There were 7 patients who underwent renal biopsies, with pathologic diagnoses of medullary cystic kidney disease or chronic tubulointerstitial damage. In addition, 6 patients had hypertension, with a duration of 0 to 6 years. Furthermore, 3 patients had overt proteinuria (>0.5 g/gCr), whereas 3 patients had normal urinary protein levels. With the inclusion of family-affected patients, the median age of patients with familial ESKD was 45 years (n = 31, 95% CI: 40–50 years) (Figure 3).

DISCUSSION

To best of our knowledge, this is the first report of a genetic analysis of *MUC1* VNTR variants detected using a combination of SRS and LRS in a Japanese cohort. We identified *MUC1* VNTR mutations in 9 patients from unrelated families. Furthermore, the detection rate of *MUC1* variants among patients with clinically suspected ADTKD was 18.8%.

Owing to the complexity of the variant hotspot, detecting *MUC1* VNTR variants by routine genetic analysis using an SRS is challenging, and other methods, such as the snapshot method or mass spectrometry, are typically required. Unfortunately, these methods are technically demanding and are only performed at a few laboratories worldwide.[5,6] Wenzel *et al.*[7] conducted SMRT sequencing and confirmed all diagnoses in European participants from 9 families with previously detected *MUC1* VNTR variants using the snapshot method. Wang *et al.*[8] also revealed the usefulness of SMRT sequencing for detecting *MUC1* VNTR mutations in a large Chinese family. Nevertheless, owing to their study design, these 2 studies did not determine the detection rate of *MUC1* VNTR variants in patients with clinically suspected ADTKD. In this study, we clarified that the diagnostic rate of *MUC1* variants in patients with clinically suspected ADTKD was 18.8%, which may facilitate more accurate probabilities, especially during genetic counseling.

Data on the prevalence of ADTKD-*MUC1* are limited, although previous studies have revealed a prevalence of 0.7 to 4 per million in the United States and Ireland.[3,13] A study in England revealed an estimated prevalence of ADTKD-*UMOD* of 9 per million, suggesting ADTKD-*UMOD* is the most frequent non-polycystic kidney disease.[14] In our study, 9 patients

Table 2. Clinicopathologic characteristics of patients with *MUC1* variant

ID	Modality/tag no. for LRS	Age	Sex	FH	Renal biopsy	Pathologic diagnosis/findings	HTN (age at diagnosis)	HUA (age at diagnosis)	Childhood anemia (age at diagnosis)	Serum potassium (mEq/l)	eGFR	UP	Inserted nucleotide	Repeat number of unit inserted a nucleotide
SC359	SRS	39	M	+	+	MCKD	+ (36 yr-)	+ (34 yr-)	–	4.1	40	1.6	C	2
SC356	LRS/bc1003	42	F	+	+	MCKD	+ (41 yr-)	–	–	4.9	31.9	0.07	C	10
SC370	LRS/bc1004	48	M	+	–	–	–	NA	–	4.1	RRT (32 yr-)	NA	G	5
SC416	LRS/bc1006	33	M	+	+	Chronic TIN	–	+ (29 yr-)	–	4.6	23	1.6	C	6
SC489	LRS/bc1007	54	M	+	+	Chronic damage of interstitium	+ (54 yr-)	+ (53 yr-)	–	5.6	6.5	1.4	C	7
SC512	LRS/bc1028	37	M	–	+	MCKD	–	+ (NA)	–	4.8	16.4	0.13	C	14
SC534	LRS/bc1009	41	F	+	+	MCKD	+ (42 yr-)	–	–	3.9	RRT (40 yr-)	0.27	C	19
SC566	LRS/bc1032	50	M	–	+	MCKD	+ (42 yr-)	–	–	3.8	10	0.27	C	40
SC656	LRS/bc1018	46	F	+	–	–	+ (40 yr-)	+ (NA)	–	4.0	41	0.12	C	7

CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate (ml/min per 1.73 m^2); F, female; FH, family history of CKD; HTN, hypertension; HUA, hyperuricemia; ID, identification; LRS, long-read sequencer; M, male; MCKD, medullary cystic kidney disease; NA, not available; RRT; renal replacement therapy; SRS, short-read sequencer; TIN, tubulointerstitial nephritis; UP, urinary protein (g/gCr).

Table 3. Details of the variants detected in the *MUC1* VNTR region

ID	Modality/tag no. for LRS	Inserted nucleotide	Repeat number of VNTR unit with nucleotide insertion	Sequence of 60-nucleotide repeat unit harboring a nucleotide insertion
SC359	SRS	C	2	GCC CAC GGT GTC ACC TCG GCC CCG GAC ACC AGG CCG GCC CCG GGC TCC ACC GCC CCC CCCA
SC416	LRS/bc1006	C	6	
SC489	LRS/bc1007	C	7	
SC534	LRS/bc1009	C	19	
SC656	LRS/bc1018	C	7	
SC512	LRS/bc1028	C	14	
SC566	LRS/bc1032	C	40	
SC356	LRS/bc1003	C	10	GCC CAC GGT GTC ACC TCG GCC CCG GAG AGC AGG CCG GCC CCG GGC TCC ACC GCC CCC CCCA
SC370	LRS/bc1004	G	5	GCC CAC GGT GTC ACC TCG GCC CCG GAC ACC GAG GCC GGC CCC GGG CTC CAC CGC CCC CCAA

ID, identification; LRS, long-read sequencer; SRS, short-read sequencer; VNTR, variable number tandem repeat.

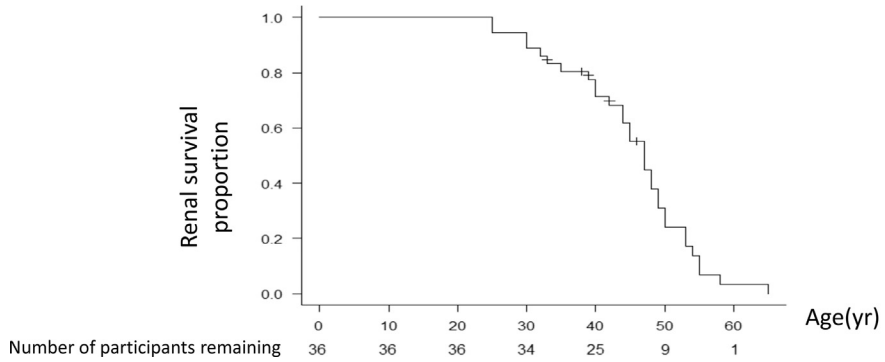

Figure 3. Probability of developing ESKD. The median age for developing ESKD was 45 years, including family-affected patients (n = 31). ESKD, end-stage kidney disease.

had *MUC1* variants, whereas 6 had *UMOD* variants, suggesting ADTKD-*MUC1* could be 1 of the most common disease subtypes of ADTKD.

Despite the structural complexity of the *MUC1* VNTR region, SRS could detect variants in 3 patients, all of whom had the same mutation: a cytosine inserted into 7 consecutive cytosines at the end of the 60-nucleotide unit in the second repeat of the VNTR. It was presumed that variants in the shallow part of the *MUC1* VNTR (until the third repeat of the VNTR) would be detectable even though the DNA was fragmented (100–200 base pair per read) as part of the sample preparation for SRS analysis. Yamamoto *et al.*[15] identified another frameshift mutation located before the *MUC1* VNTR region using an SRS (whole exome sequence). This mutation also generates an abnormally truncated protein in the same manner as would the mutation positioned within the VNTR.[15] These results suggest that SRS analysis may be useful for detecting variants positioned before or within the shallow region of the VNTR.

In our study, 7 of the 9 patients with *MUC1* VNTR variants had the same 60-nucleotide unit sequence harboring the inserted cytosine (Table 2). In previous studies in which long-read sequencing genetic analysis for *MUC1* VNTR was performed, the same variant was detected in 9 families in a European cohort[7] and a large Chinese family.[8] In addition, this mutation has been detected using the snapshot method and was defined as "27dupC."[12] Olinger *et al.*[16] reported that the prevalence of the 27dupC variant was 93.5% in 2 registries of ADTKD in Europe and the United States. In accordance with these findings, our data indicated that the 27dupC variant was the most common *MUC1* mutation in a Japanese cohort.

The number of VNTR repeats has been reported to be 20 to 25[3]; however, in our study, the VNTR repeat number ranged from 30 to 82, with 80% of the alleles having 30 to 36 repeats. A previous study in Europe reported that the percentage of alleles with 30 to 36 repeats was 39%, which may be underpinned by the ethnic homogeneity of the Japanese population.

The present study revealed that the median age at developing ESKD was 45 years (95% CI: 40–50 years) in patients with ADTKD-*MUC1*, including familial-affected patients in a Japanese cohort (9 families, $n = 31$). A previous study evaluated 147 individuals, including patients with *MUC1* mutations and their affected family members, and revealed that the mean age at developing ESKD was 44.9 ± 15.4 years.[17] In contrast, recent research using different cohorts indicated that the median time of kidney survival in 104 patients with ADTKD-*MUC1* was 36 years (interquartile range: 30–46 years), which may be explained by the exclusion of affected family members who did not undergo genetic analysis.[16] In addition, interfamilial and intrafamilial variations have been reported in patients with *MUC1* variants,[8,17,18] although the existence of modifier genes or other factors remains unknown.

In the same study period, we also detected mutations in *HNF1B*, a causative gene for ADTKD, in 19 patients. Nevertheless, patients with clinically suspected *HNF1B* mutations were excluded from our study at registration, because most of these patients were clinically distinguishable from those with ADTKD-*MUC1* because of morphologic abnormalities in the renal-urinary tract system and/or an autosomal-dominant familial history of early onset diabetes and/or hypomagnesemia, and we have previously reported on this patient population.[19] Indeed, a Kidney Disease: Improving Global Outcomes consensus report and recent review of ADTKD suggested that the term ADTKD-*HNF1B* should only be reserved for those cases in which kidney tubulointerstitial fibrosis is the leading manifestation because only a few cases present with tubulointerstitial disease only.[1,20] Childhood anemia, mild hypotension, and mild hyperkalemia are thought to be specific features of ADTKD-REN.

Nevertheless, a recent international cohort study on clinical characteristics of 111 ADTKD-REN patients reported that approximately 70% of the patients presented to medical institutions for chronic kidney disease or gout.[21] According to this report, childhood anemia was present in 75.8% of the patients but its severity was relatively mild; mean hemoglobin levels were 9.6, 10.1, and 10.5 g/dl for ages <10 years, 10 to <15 years, and 15 to <20 years, respectively.[21] Although hypotension was not reported in the article, the severity of hyperkalemia was mild and the mean serum potassium level in patients who were not taking fludrocortisone was 4.8 mEq/l.[21] Childhood anemia, hypotension, or hyperkalemia may often be overlooked, and they are not always specific to ADRKD-REN; thus, we did not exclude participants with these findings. ADTKD-*SEC61A1* is much more rare than other ADTKD subtypes, with 6 families reported as of date.[22] Patients harboring *SEC61A1* mutation present with specific features, such as intrauterine growth retardation, cleft palate, congenital anemia, neutropenia, and immunodeficiency.[3,22] Although the clinical manifestations of ADTKD-REN have not been completely clarified owing to the small number of reported cases, patients presenting with these specific features can be clearly distinguished from those with ADTKD-*MUC1*.

Recent studies on the molecular and cellular mechanisms of ADTKD-*MUC1* have revealed that this condition is a toxic proteinopathy caused by the intracellular accumulation of misfolded MUC1 protein.[11,23] Dvela-Levitt *et al.*[24] reported the use of a small molecule, BRD4780, to reroute the secretory pathway to lysosomes, revealing MUC1fs were eliminated in knock-in mice and patient organoids and highlighting the therapeutic potential of BRD4780. Accordingly, an accurate diagnosis of *MUC1* variants may be crucial for the development of effective therapeutic agents.

Our study had several limitations. First, all of our study participants were Japanese, and our sample size was small owing to the short study duration. Given the retrospective nature of this study, we were also unable to obtain longitudinal clinical information, such as renal prognosis. Second, single-base insertions were observed even in alleles without mutations. Owing to high error rate, it was challenging to achieve high-quality assembly to detect single-nucleotide variants or indels using an LRS. It was necessary to combine sequencing technologies to detect all the different types of genetic variations, which increases the cost and complexity of projects. Thus, PacBio devised a new sequencing system, named circular consensus sequencing. Circular consensus sequencing produces high-fidelity reads with 99.8% accuracy and an average length of 13.5 kilobase, which drastically reduces the sequencing error rate.[25] This sequencing method was also applied in our study. Hence, considering improvements in the sequencing accuracy, single-base insertions were possibly caused by PCR error. Even though KOD Hot Start DNA polymerase was optimized for the amplification of the most difficult targets,[26] other DNA polymerases, the fidelity of which is higher than that of KOD (e.g., Phusion), may contribute in reducing the PCR error rate.[27] Recent studies have found an amplification-free protocol for targeted enrichment using the clustered regularly interspaced short palindromic repeats/Cas9 system.[28,29] This new method would likely provide better results. Third, we have not validated an alternative method that could facilitate a diagnosis, such as immunostaining of kidney specimens or urine-derived cells. Kirby *et al.*,[5] who first reported the causative mutation for ADTKD-*MUC1*, performed an immunohistochemical analysis of kidney samples from patients with *MUC1* mutations. They reported that, although immunostaining in normal controls resulted in nonspecific staining, patients with *MUC1* mutations exhibited a specific intracellular staining pattern in the loop of Henle, distal tubules, and collecting ducts.[5] Similar immunostaining analyses of the mutant MUC1 protein have been reported.[11,13,18,30] Yamamoto *et al.*[15] detected mutant MUC1 protein in urine exosomes. Although the clinical manifestations of ADTKD-*MUC1* are nonspecific, these nongenetic analyses may support the clinical diagnosis. Thus, it is necessary to verify the usefulness of these nongenetic analyses in conjunction with the results of our study.

In conclusion, our study reveals that combination genetic analysis with SRS and LRS is useful for detecting *MUC1* VNTR variants in a Japanese cohort and that the prevalence of ADTKD-*MUC1* may be high in this population. Combination genetic analysis using SRS and LRS is likely to improve the diagnosis rate of ADTKD-*MUC1* and may contribute to the development of optimal treatment approaches.

DISCLOSURE

TH reports receiving research funding from Otsuka Pharmaceutical Co., Ltd.

ACKNOWLEDGMENTS

The authors thank all the study participants, their social guardians, and their primary doctors. The authors are profoundly grateful to Ms. Yoshimi Nozu, Ms. Ming Juan Ye, and Ms. Yuko Noguchi (Kobe University) for their excellent technical assistance. The authors thank Editage (www.editage.com) for English language editing.

This work was supported by the Japan Society for the Promotion of Science (to NM) under grant number 18K08243.

DATA STATEMENT

The data are not available for public access owing to patient privacy concerns but are available from the corresponding author on reasonable request.

SUPPLEMENTARY MATERIAL

Supplementary File (PDF)

Table S1. The gene list constructed using HaloPlex (version 2, 128 genes).

Table S2. The gene list constructed using HaloPlex (version 4, 172 genes).

Table S3. The gene list constructed using HaloPlex (version 5, 159 genes).

Table S4. The gene list constructed using HaloPlex (version 6, 164 genes).

Table S5. The gene list constructed using HaloPlex (version 7, 181 genes).

Table S6. The gene list constructed using HaloPlex (version 8, 183 genes).

Table S7. Version of HaloPlex used for each patient analyzed by SRS.

Table S8. The identified sequences of VNTR region for each patient analyzed by LRS.

Table S9. Details of the variants detected by SRS.

Table S10. Clinicopathologic characteristics of patients with SRS-detected variants other than *MUC1*.

Table S11. Summary of identifying alleles of each patient analyzed by LRS.

Figure S1. Distribution of repeat number of VNTR of each allele.

REFERENCES

1. Eckardt KU, Alper SL, Antignac C, et al. Autosomal dominant tubulointerstitial kidney disease: diagnosis, classification, and management—a KDIGO consensus report. *Kidney Int*. 2015;88:676–683. https://doi.org/10.1038/ki.2015.28
2. Bolar NA, Golzio C, Živná M, et al. Heterozygous loss-of-function SEC61A1 mutations cause autosomal-dominant tubulo-interstitial and glomerulocystic kidney disease with anemia. *Am J Hum Genet*. 2016;99:174–187. https://doi.org/10.1016/j.ajhg.2016.05.028
3. Devuyst O, Olinger E, Weber S, et al. Autosomal dominant tubulointerstitial kidney disease. *Nat Rev Dis Primers*. 2019;5:60. https://doi.org/10.1038/s41572-019-0109-9
4. Patton S, Gendler SJ, Spicer AP. The epithelial mucin, MUC1, of milk, mammary gland and other tissues. *Biochim Biophys Acta*. 1995;1241:407–423. https://doi.org/10.1016/0304-4157(95)00014-3
5. Kirby A, Gnirke A, Jaffe DB, et al. Mutations causing medullary cystic kidney disease type 1 lie in a large VNTR in MUC1 missed by massively parallel sequencing. *Nat Genet*. 2013;45:299–303. https://doi.org/10.1038/ng.2543
6. Ekici AB, Hackenbeck T, Morinière V, et al. Renal fibrosis is the common feature of autosomal dominant tubulointerstitial kidney diseases caused by mutations in mucin 1 or uromodulin. *Kidney Int*. 2014;86:589–599. https://doi.org/10.1038/ki.2014.72
7. Wenzel A, Altmueller J, Ekici AB, et al. Single molecule real time sequencing in ADTKD-MUC1 allows complete assembly of the VNTR and exact positioning of causative mutations. *Sci Rep*. 2018;8:4170. https://doi.org/10.1038/s41598-018-22428-0
8. Wang GQ, Rui HL, Dong HR, et al. SMRT sequencing revealed to be an effective method for ADTKD-MUC1 diagnosis through follow-up analysis of a Chinese family. *Sci Rep*. 2020;10:8616. https://doi.org/10.1038/s41598-020-65491-2
9. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17:405–424. https://doi.org/10.1038/gim.2015.30
10. Kanda Y. Investigation of the freely available easy-to-use software "EZR" for medical statistics. *Bone Marrow Transplant*. 2013;48:452–458. https://doi.org/10.1038/bmt.2012.244
11. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–26. https://doi.org/10.1038/nbt.1754
12. Živná M, Kidd K, Přistoupilová A, et al. Noninvasive immunohistochemical diagnosis and novel MUC1 mutations causing autosomal dominant tubulointerstitial kidney disease [published correction appears in *J Am Soc Nephrol*. 2020;31:892]. *J Am Soc Nephrol*. 2018;29:2418–2431. https://doi.org/10.1681/ASN.2018020180
13. Cormican S, Connaughton DM, Kennedy C, et al. Autosomal dominant tubulointerstitial kidney disease (ADTKD) in Ireland. *Ren Fail*. 2019;41:832–841. https://doi.org/10.1080/0886022X.2019.1655452
14. Gast C, Marinaki A, Arenas-Hernandez M, et al. Autosomal dominant tubulointerstitial kidney disease—UMOD is the most frequent non polycystic genetic kidney disease. *BMC Nephrol*. 2018;19:301. https://doi.org/10.1186/s12882-018-1107-y
15. Yamamoto S, Kaimori J-Y, Yoshimura T, et al. Analysis of an ADTKD family with a novel frameshift mutation in MUC1 reveals characteristic features of mutant MUC1 protein. *Nephrol Dial Transplant*. 2017;32:2010–2017. https://doi.org/10.1093/ndt/gfx083
16. Olinger E, Hofmann P, Kidd K, et al. Clinical and genetic spectra of autosomal dominant tubulointerstitial kidney disease due to mutations in UMOD and MUC1. *Kidney Int*. 2020;98:717–731. https://doi.org/10.1016/j.kint.2020.04.038
17. Bleyer AJ, Kmoch S, Antignac C, et al. Variable clinical presentation of an MUC1 mutation causing medullary cystic kidney disease type 1. *Clin J Am Soc Nephrol*. 2014;9:527–535. https://doi.org/10.2215/CJN.06380613
18. Ayasreh N, Bullich G, Miquel R, et al. Autosomal dominant tubulointerstitial kidney disease: clinical presentation of patients with ADTKD-UMOD and ADTKD-MUC1. *Am J Kidney Dis*. 2018;72:411–418. https://doi.org/10.1053/j.ajkd.2018.03.019

19. Nagano C, Morisada N, Nozu K, et al. Clinical characteristics of HNF1B-related disorders in a Japanese population. *Clin Exp Nephrol*. 2019;23:1119–1129. https://doi.org/10.1007/s10157-019-01747-0

20. Mabillard H, Sayer JA, Olinger E. Clinical and genetic spectra of autosomal dominant tubulointerstitial kidney disease. *Nephrol Dial Transplant*. Published online September 14, 2021. https://doi.org/10.1093/ndt/gfab268.

21. Živná M, Kidd K, Zaidan M, et al. An international cohort study of autosomal dominant tubulointerstitial kidney disease due to REN mutations identifies distinct clinical subtypes. *Kidney Int*. 2020;98:1589–1604. https://doi.org/10.1016/j.kint.2020.06.041

22. Bleyer AJ, Wolf MT, Kidd KO, Zivna M, Kmoch S. Autosomal dominant tubulointerstitial kidney disease: more than just HNF1β. *Pediatr Nephrol*. Published online May 22, 2021. https://doi.org/10.1007/s00467-021-05118-4.

23. Knaup KX, Hackenbeck T, Popp B, et al. Biallelic expression of mucin-1 in autosomal dominant tubulointerstitial kidney disease: implications for nongenetic disease recognition. *J Am Soc Nephrol*. 2018;29:2298–2309. https://doi.org/10.1681/ASN.2018030245

24. Dvela-Levitt M, Kost-Alimova M, Emani M, et al. Small molecule targets TMED9 and promotes lysosomal degradation to reverse proteinopathy. *Cell*. 2019;178:521–535.e23. https://doi.org/10.1016/j.cell.2019.07.002

25. Wenger AM, Peluso P, Rowell WJ, et al. Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. *Nat Biotechnol*. 2019;37:1155–1162. https://doi.org/10.1038/s41587-019-0217-9

26. Terpe K. Overview of thermostable DNA polymerases for classical PCR applications: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*. 2013;97:10243–10254. https://doi.org/10.1007/s00253-013-5290-2

27. McInerney P, Adams P, Hadi MZ. Error rate comparison during polymerase chain reaction by DNA polymerase. *Mol Biol Int*. 2014;2014:287430. https://doi.org/10.1155/2014/287430

28. Ebbert MTW, Farrugia SL, Sens JP, et al. Long-read sequencing across the C9orf72 "GGGGCC" repeat expansion: implications for clinical use and genetic discovery efforts in human disease. *Mol Neurodegener*. 2018;13:46. https://doi.org/10.1186/s13024-018-0274-4

29. Höijer I, Tsai YC, Clark TA, et al. Detailed analysis of HTT repeat elements in human blood using targeted amplification-free long-read sequencing. *Hum Mutat*. 2018;39:1262–1272. https://doi.org/10.1002/humu.23580

30. Yu SMW, Bleyer AJ, Anis K, et al. Autosomal dominant tubulointerstitial kidney disease due to MUC1 mutation. *Am J Kidney Dis*. 2018;71:495–500. https://doi.org/10.1053/j.ajkd.2017.08.024