www.nature.com/jhg

ARTICLE

Intronic LINE-1 insertion in *SLCO1B3* as a highly prevalent cause of rotor syndrome in East Asian population

Young-gon Kim[1], Hobin Sung[1], Ho Seob Shin[1], Man Jin Kim[1], Jee-Soo Lee[1], Sung-Sup Park[1] and Moon-Woo Seong[1]✉

Rotor syndrome is caused by digenic loss-of-function variants in *SLCO1B1* and *SLCO1B3* but only a few studies have reported co-occurring inactivating variants from both genes. A rotor syndrome-causing long interspersed element-1 (LINE-1) insertion in *SLCO1B3* had been reported to be highly prevalent in the Japanese population but there has been no additional report. In spite of its known association with various human diseases, LINE-1 is hard to detect with current sequencing technologies. In this study, we aimed to devise a method to screen the LINE-1 insertion variant and investigate the frequency of this variant in various populations. A chimeric sequence, that was generated by concatenating the reference sequence at the junction and a part of inserted LINE-1 sequence, was searched from 725 raw sequencing data files. In cases containing the chimeric sequence, confirmatory long-range PCR and gap-PCR were performed. In total, 95 (13.1%) of 725 patients were positive for the chimeric sequence, and all were confirmed to have the *SLCO1B3* LINE-1 insertion by PCR-based tests. The same chimeric sequence was searched from the 1000 Genomes Project data repository and the carrier frequency was remarkably high in the East Asian populations (10.1%), especially in Southern Han Chinese (18.5%), but almost absent in other populations. This *SLCO1B3* LINE-1 insertion should be screened in a population-specific manner under suspicion of Rotor syndrome and the methods proposed in this study would enable this in a simple way.

Journal of Human Genetics; https://doi.org/10.1038/s10038-021-00967-1

INTRODUCTION

Long-interspersed element-1 (LINE-1) retrotransposons are one of the most abundant classes of mobile DNAs, accounting for 17% of the human genome [1, 2]. While the vast majority of the LINE-1 sequences in the human genome are inactive due to rearrangements and truncations [1, 3, 4], a small subset is thought to remain active [5, 6]. When out of control, LINE-1 retrotransposition can lead to human diseases such as cancer and heritable genetic diseases [7, 8].

Rotor syndrome is a rare disorder that presents with chronic, fluctuating, and benign conjugated hyperbilirubinemia. Although it has long been known as a familial disease since its discovery in 1948 [9], it was not until 2012 that the pathogenesis of this disease was understood at the molecular level [10]. It is now established that Rotor syndrome is caused by the complete deficiency of organic anion transporting polypeptides OATP1B1 and OATP1B3, which mediate reuptake of conjugated bilirubin into the liver; these two proteins are encoded by the genes *SLCO1B1* and *SLCO1B3*, respectively, and the molecular diagnosis of Rotor syndrome is made by the identification of biallelic inactivating mutations from both genes [11]. The two genes are located close to each other in chromosome 12p, and three haplotypes containing pathogenic variants from both *SLCO1B1* and *SLCO1B3* were reported by van de Steeg et al. [10].

In a study performed in Japan by Kagawa et al. [12], genomic analysis of *SLCO1B1* and *SLCO1B3* was performed to determine the genetic causes of seven Rotor syndrome patients, comprising six Japanese and one Filipino. All six Japanese patients were homozygous for the novel haplotype discovered in their study. The haplotype consisted of a previously known nonsense variant of *SLCO1B1*, c.1738C>T, p.R580*, and a novel LINE-1 insertion in intron 6 of *SLCO1B3* (ClinVar ID: 977762, designated as intron 5 in Kagawa et al. [12]). In their study, the 6.1kbp LINE-1 insertion located at the 3′ flanking region of exon 6 was detected by long-range PCR and subsequent sequencing by primer walking. The loss-of-function effect of this insertion via aberrant splicing was proven by RNA analysis and immunohistochemistry of liver biopsy tissues. An investigation of 554 healthy volunteers revealed that this insertion is not rare in the Japanese population (estimated allele frequency 5.4%).

In this study, we aimed to determine the population-specific frequency of the *SLCO1B3* intron 6 LINE-1 insertions. We performed LINE-1 chimeric sequence analysis from both the Seoul National University Hospital (SNUH) repository and the public repository of the 1000 Genomes Project [13]. Positive results from chimeric sequence analysis of SNUH repository were confirmed by long-range PCR and/or gap-PCR. Based on the remarkably high frequencies of this insertion found in specific populations, we suggest a population-specific approach for the genetic testing of suspected Rotor syndrome patients.

METHODS

Participants

This study was approved by the Institutional Review Board of SNUH. Informed consent for genetic testing was obtained from all included

[1]Department of Laboratory Medicine, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Republic of Korea. ✉email: mwseong@snu.ac.kr

Received: 11 June 2021 Revised: 22 July 2021 Accepted: 25 July 2021
Published online: 05 August 2021

participants or their guardians in the case of minor patients. First, patients who had been requested for the cholestasis gene panel were targeted. Among the 91 patients who were tested for the cholestasis gene panel in SNUH from October 2017 to November 2020, four patients with biallelic pathogenic variants in *SLCO1B1* were further analyzed for the detection of *SLCO1B3* LINE-1 insertion (Table 1).

To estimate the allele frequency of the *SLCO1B3* intron 6 LINE-1 insertions in the general Korean population, the raw sequencing data (BAM files) of 749 patients who were tested for all types of NGS panel tests in SNUH from January 2020 to November 2020 were screened. To simulate the normal population, in terms of hyperbilirubinemia, we excluded patients who requested the cholestasis gene panel, leaving 727 patients to be included in the analysis. Among them, the carrier and allele frequencies were derived from 725 Korean patients.

To estimate the frequency of *SLCO1B3* intron 6 LINE-1 insertion in various populations, raw sequencing data from the 1000 Genomes Project phase 3, which consists of BAM files from 2535 individuals from 26 populations, were utilized. For each individual, both the low-coverage whole-genome sequencing (WGS) data and exome sequencing data were used, resulting in a total of 5070 data files (Supplementary Table S1).

NGS panel test

In SNUH, gene panel tests are performed as virtual panels in which a predefined set of genes is analyzed after sequencing the whole set of target genes from all panels. Therefore, the frequency estimation of the *SLCO1B3* intron 6 LINE-1 insertion was possible by analyzing the repository of raw data from gene panel tests that did not include *SLCO1B3* in their target gene lists. The DNA capture probes for 1836 genes that included the target genes of 67 panels were designed using the Agilent SureDesign web-based application (https://earray.chem.agilent.com/suredesign/, Agilent Technologies, Santa Clara, CA, USA). The target regions included protein-coding exons with 25 bp intron flanking regions and deep-intronic regions of disease-causing sequence variants reported in the Human Gene Mutation Database [14]. DNA was extracted using a Chemagic 360 instrument (Perkin Elmer, Baesweiler, Germany) and sheared using a Covaris E220 focused-ultrasonicator (Covaris, Woburn, MA, USA). We used 500 ng of total input genomic DNA. Library preparation was performed according to the Agilent SureSelectXT Target Enrichment protocol (Agilent Technologies). Paired-end 150-bp sequencing was performed using the NextSeq 550 Dx platform (Illumina, San Diego, CA, USA) and Nextgene Version 2.4.0.1 (Softgenetics, State College, PA, USA) was used for bioinformatic processes from the alignment to the annotation. NM_006446.4 and NM_019844.3 were used as reference transcripts for *SLCO1B1* and *SLCO1B3*, respectively.

Analysis of raw sequencing data files

In search of the *SLCO1B3* LINE-1 insertion from the patients described in Table 1, we manually reviewed raw sequencing data (BAM files) using Integrative Genomics Viewer (IGV, Version 2.8.10) [15]. After identifying and confirming the LINE-1 insertion from those patients, for its frequency estimation, we searched the repository of BAM files following the steps described below. First, we created sampled BAM files containing only reads that were mapped to the genetic location of LINE-1 insertion, which is chr12:21014093 (GRCh 37), by specifying this location to the SAMtools software (Version 1.9) [16].

The command used for this operation was:

samtools view -bh $file 12:21014093-21014093 > "${file%.bam}_extracted.bam"

For the data stored in the 1000 Genomes Project repository (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/phase3, last accessed 21 May 2021), the same command was used to download the sampled BAM file by replacing "$file" with the absolute path of the file in the repository.

After converting the sampled BAM files to text files using SAMtools, we searched the chimeric sequence, which was made by concatenating the reference sequence and a part of inserted LINE-1 sequence at the junction, from the text files using the grep tool provided in CentOS 7. The chimeric sequence we used was GGTAAGAATTAATAGTGACAGTGAGGAACAGCTCC, which consisted of the first 22 nucleotides from the reference sequence and the latter 13 nucleotides from the inserted LINE-1 sequence.

The commands we used for these operations were:

samtools view $file > $file".txt"

find *.bam.txt -exec grep -l GGTAAGAATTAATAGTGACAGTGAGGAACAGCTCC {} \;

Table 1. Characteristics of patients previously detected with pathogenic *SLCO1B1* variants from cholestasis gene panel tests and the results of *SLCO1B3* LINE-1 insertion tests from these patients

Subject	Ethnicity	Age	Gender	TB/DB (mg/dl)	*SLCO1B1* variant	*SLCO1B3* LINE-1 insertion		
						Estimated VAF from BAM files	Long-range PCR	Gap-PCR
RS1	Korean	17	F	4/2.7	c.1738C>T, p.R580*, homozygote	82.5% (33/40)[a]	Homozygous	+
RS2	Korean	2	M	2.3/1.9	c.1738C>T, p.R580*, homozygote	74.1% (60/81)	Homozygous	+
RS3	Korean	14	F	5.2/3.4	c.1738C>T, p.R580*, homozygote	90.0% (45/50)	Homozygous	+
RS4	Korean	9	M	6.1/4.8	c.1738C>T, p.R580*, homozygote	85.2% (98/115)	Homozygous	+

TB total bilirubin, *DB* direct bilirubin, *BAM* binary alignment map, *VAF* variant allele frequency, *LINE-1* long interspersed nuclear elements-1
[a]Right-clipped reads/read depth

The variant allele frequency (VAF) of the LINE-1 insertion was estimated from the sampled BAM files by dividing the number of right-clipped reads by the total read depth.

For the search of individuals harboring *SLCO1B1* c.1738C>T, p.R580* variant from 1000 Genomes Project VCF (Variant Call Format) repository, Data Slicer software (http://grch37.ensembl.org/Homo_sapiens/Tools/DataSlicer?db=core, last accessed, 21 May 2021) was firstly used to extract the variants in the target genomic region from VCF files and carriers of the variant were searched from Microsoft Excel.

Confirmatory tests—long-range PCR and Gap-PCR followed by Sanger sequencing

To confirm the existence of *SLCO1B3* intron 6 LINE-1 insertion, PCR tests were performed using residual DNA from the initial gene panel tests. The previously described primers were used for long-range PCR and gap-PCR (Supplementary Table S2, Fig. 1). The expected band size of long-range PCR was 6.5 kbp for the LINE-1 insertion-positive alleles and 416 bp for insertion-negative alleles. Gap-PCR targeted both junctions of the LINE-1 insertion and was expected to produce a 349 bp amplicon from the 5′ junction and 800 bp amplicon from the 3′ junction, respectively, from LINE-1 insertion-positive samples.

For the four patients in Table 1, who had biallelic *SLCO1B1* variants and features of homozygous *SLCO1B3* LINE-1 insertion from their BAM files, we performed long-range PCR and gap-PCR followed by Sanger sequencing to confirm the insertion. For 95 patients who screened positive from the search for chimeric sequences from the repository of BAM files, long-range PCR was initially performed. Cases with the 6.5 kbp band from the long-range PCR were considered to have LINE-1 insertion. For patients whose long-range PCR results were negative, gap-PCR was additionally performed.

In long-range PCR, the cycling parameters included an initial denaturation at 94 °C for 2 min; 35 cycles of 94 °C for 15 s, 65 °C for 30 s, and 68 °C for 8 min; and a final extension at 68 °C for 7 min. In gap-PCR, cycle parameters included an initial denaturation at 94 °C for 5 min; 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min; and a final extension at 72 °C for 7 min. Sanger sequencing was performed using an ABI 3730 analyzer (Thermo Fisher Scientific, MA, USA), and SnackVar software [17] was used for the analysis of trace files.

RESULTS

Identification of the *SLCO1B3* intron 6 LINE-1 insertions from patients previously detected to have pathogenic *SLCO1B1* variants

All four patients described in Table 1 were found to have BAM file findings suggestive of homozygous LINE-1 insertion near the exon 6–intron 6 junction of *SLCO1B3* (Table 1, Fig. 2). Most of the reads covered at the genomic location of LINE-1 insertion, which is 21 bp downstream of the exon 6–intron 6 junction, were soft-clipped on their right side. The right-clipped sequence was identical in all clipped reads starting with "GAGGAACAGCTCC" and could not be mapped to other genomic regions in the reference genome. A sudden drop in the read depth was observed in the homozygous LINE-1 insertion carriers.

Long-range PCR and gap-PCR followed by Sanger sequencing confirmed homozygous LINE-1 insertion in all four patients (Table 1 and Fig. 3). The 6.5 kbp band, which was observed in the study by Kagawa et al. [12], was detected in all four patients. The sequences of both ends of the LINE-1 insertion revealed by Sanger sequencing of gap-PCR products were also identical to the sequence detected by Kagawa et al.

The estimated VAF of confirmed homozygous LINE-1 insertions, which was derived from the number of right-clipped reads and the total read depth, ranged from 74.1% to 90.0%, suggesting the tendency towards underestimation of actual VAF by this method.

Frequency estimation of the *SLCO1B3* intron 6 LINE-1 insertions from SNUH repository

Among the 725 BAM files of Korean patients included from the SNUH data repository, 95 (13.1%) were positive for the target chimeric sequence (Supplementary Tables S3 and 2). The chimeric sequence was negative for the two BAM files from non-Korean patients. Among 95 patients who were positive for the chimeric sequence, 91 were confirmed to carry the *SLCO1B3* intron 6 LINE-1 insertion by long-range PCR (4 homozygous and 87 heterozygous). Long-range PCR negative specimens were subjected to gap-PCR and the results were positive in all four patients with expected band sizes (CR07, CR10, CR13, and CR34). Based on the BAM file findings and gap-PCR results, we assumed that these four patients were positive for the heterozygous *SLCO1B3* intron 6 LINE-1 insertion. We reasoned that the negative result from long-range PCR is due to DNA degradation during storage, considering that long-range PCR requires high-quality DNA for successful amplification. The mean lead time from the initial NGS test to the PCR test was 6.7 months for 91 long-range PCR positive samples and 9.5 months for the four long-range PCR negative samples.

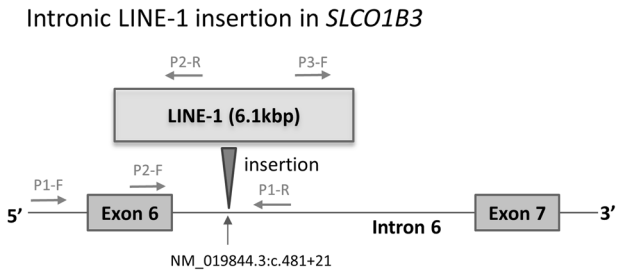

Fig. 1 Long-range PCR was performed using primers P1-F and P1-R. The expected size of the amplicon was 6.5 kbp for the alleles with LINE-1 insertion and 0.4 kbp for the insertion negative alleles. Gap-PCR of the 5′ junction of insertion was performed using primers P2-F and P2-R. The expected amplicon size was 349 bp for the insertion positive alleles. Gap-PCR of the 3′ junction of insertion was performed using primers P3-F and P1-R and was expected to produce an 800 bp amplicon for the insertion positive alleles. The sequences of primers are described in Supplementary Table S2

The average read depth at the genomic location of LINE-1 insertion was 68.9, 218.1, and 382.6 for the homozygous carriers, heterozygous carriers, and non-carriers, respectively. The average estimated VAF was 81.2% for homozygous carriers and 19.7% for heterozygous carriers, again suggesting the tendency of underestimation of actual VAF from the formula we used.

Among the 725 included patients, five were heterozygous carriers of the *SLCO1B1* c.1738C>T, p.R580* variant. Two of them also carried a heterozygous *SLCO1B3* intron 6 LINE-1 insertions (CR37, CR57, Supplementary Table S3), and three were negative for the insertion.

Population-specific frequencies of *SLCO1B3* intron 6 LINE-1 insertion estimated from 1000 Genomes Project data

The chimeric sequence was detected from exome BAM files of 56 individuals; among them, 25 individuals were also positive for the chimeric sequence in their low-coverage WGS BAM files (Supplementary Table S4). No individual was positive for the chimeric sequence only from the low-coverage WGS BAM files. Two of the chimeric sequence-positive individuals (HG03681 and NA19428) were considered false positives because they contained only one read with the chimeric sequence and the reads were soft-clipped on both sides, instead of right-side only, suggesting the possibility of misalignment. The estimated VAF of LINE-1 insertion in these two individuals was 3.2% and 2.8% from exome BAM files and 0% from low-coverage WGS BAM files. Except for these two, the range of estimated VAF of 54 patients was 12.7–67.7% from exome BAM files and 14.0–100.0% from low-coverage WGS BAM files. The one file with the 100.0% VAF from low-coverage WGS, NA18548, had only one read covered in the region, and the VAF from the exome BAM file was 24.0% (12/50). Excluding NA18548, the highest VAF observed for both exome and low-coverage WGS was 67.7%. Considering that the lowest estimated VAF observed from confirmed homozygous carriers in the SNUH repository was

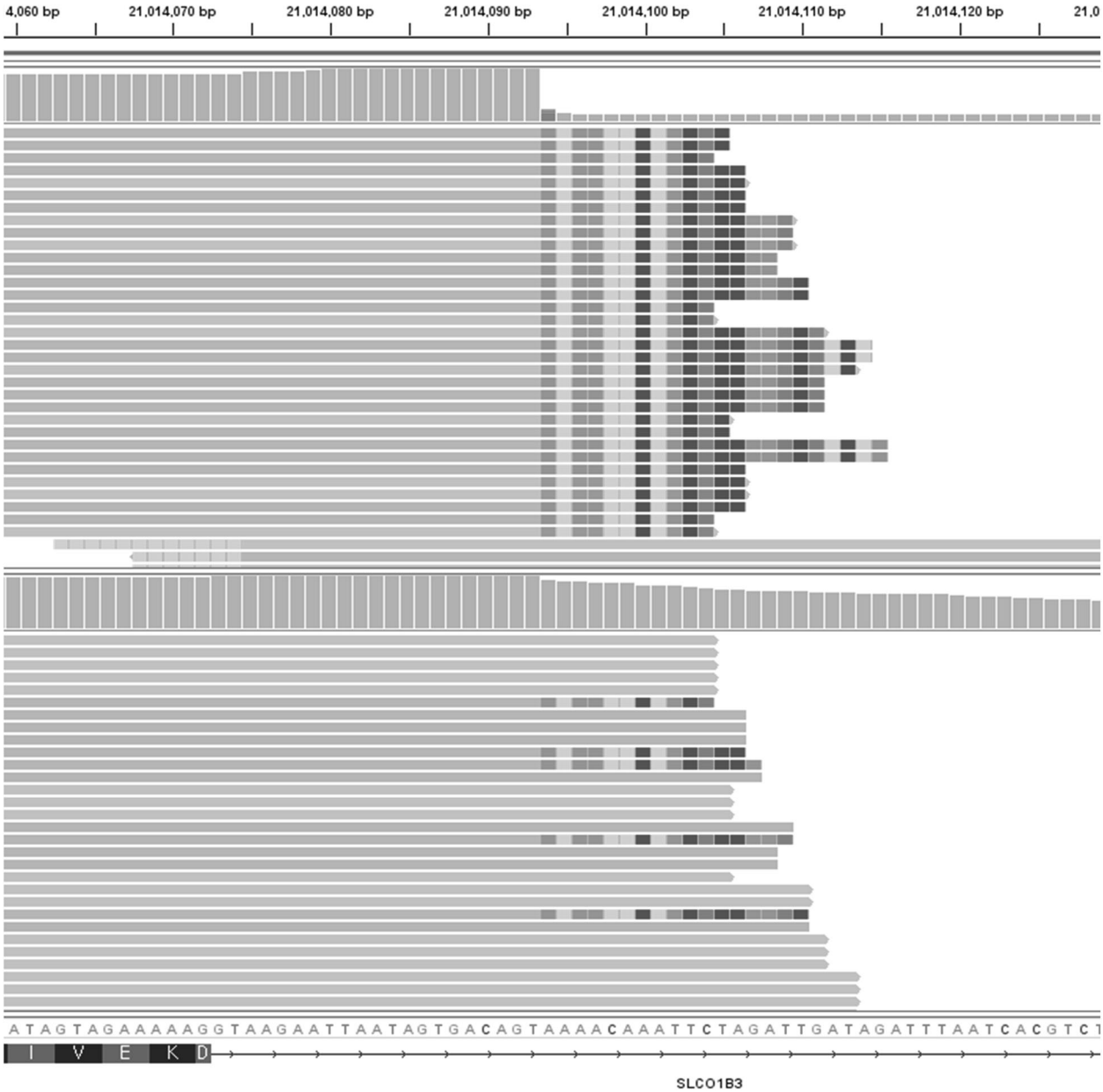

Fig. 2 BAM file findings of patients with homozygous LINE-1 insertion (upper, RS2) and heterozygous LINE-1 insertion (lower, CR2) viewed on IGV. Most of the reads covered at the junction of the insertion are soft-clipped on their right side in homozygous insertion cases. Soft-clipped reads share the same clipped sequence, which starts with "GAGGAACAGCTCC". The lower fraction of reads is soft-clipped in the heterozygous insertion case. The sudden decrease in depth is remarkable in homozygous insertion cases. Pink horizontal bars represent the reads from the forwarding strand and violet bars represent the reads from the reverse strand. BAM binary alignment map, IGV integrative genomics viewer

70.3% (CR49 from Supplementary Table S3), we regarded all 54 individuals as heterozygous carriers.

The population-specific frequencies of the *SLCO1B3* intron 6 LINE-1 insertions are summarized in Table 3. The carrier frequency was highest in the Southern Han Chinese, which was 18.5% (20/108), suggesting the founder effect of this insertion in this population. The carrier frequencies of other Chinese populations, Han Chinese, and Dai Chinese were 9.7% and 7.1%, respectively. The carrier frequency and allele frequency of the Japanese population were estimated to be 10.6% and 5.3%, respectively, and were very close to the frequencies estimated by Kagawa et al. [12], which were 10.6% and 5.4%, respectively. The overall carrier frequency in East Asia is estimated to be 10.1%. Two carriers were detected from South Asian ancestry (carrier frequency: 0.4%), and these were the only carriers detected outside of East Asian ancestry.

The results of the search for the *SLCO1B1* c.1738C>T, p.R580* variant from the 1000 Genomes Project VCF repository are presented in Supplementary Table S5. Among the 13 identified carriers, one was a homozygous carrier. Three of the heterozygous carriers were *SLCO1B3* LINE-1 insertion carriers found in this study.

DISCUSSION

In spite of its proven association to various human diseases, characterization of the LINE-1 segment from the human genome is not straightforward due to many factors, such as its abundance in the human genome, its length, which is usually about 6 kbp[1],

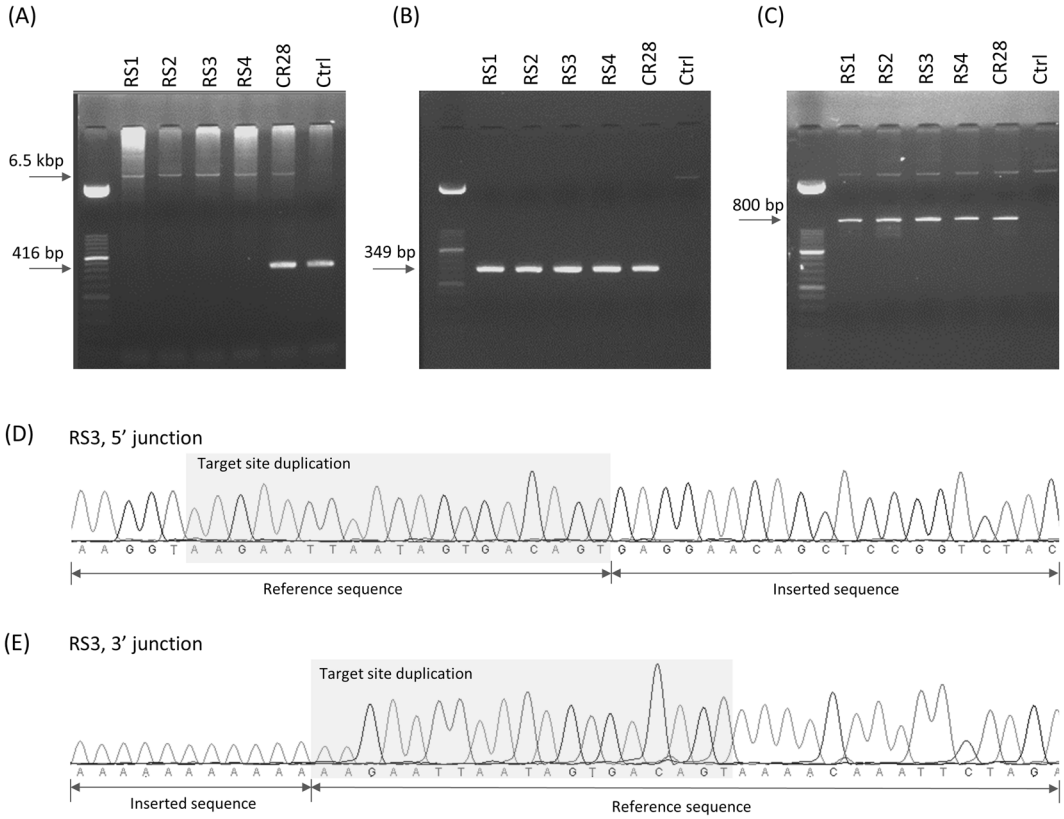

Fig. 3 Confirmatory long-range PCR and gap-PCR results of patients with homozygous intronic LINE-1 insertion (RS1-RS4), heterozygous LINE-1 insertion (CR28), and no insertion (Ctrl). **A** Gel electrophoresis of long-range PCR products. About 6.5 kbp product was detected in all patients with intronic LINE-1 insertion. In the heterozygous carrier, 416 bp product, which is the expected size from the reference genome, was additionally detected. **B**, **C** Gel electrophoresis of gap-PCR products (**B**: 5′ junction, **C**: 3′ junction). The products with expected sizes, which are 349 bp from the 5′ junction and 800 bp from 3′ junction, were detected in all the homozygous patients and heterozygous carriers. **D**, **E** The results of Sanger sequencing performed on the gap-PCR products (from RS3). The same target site duplication sequence and the LINE-1 insertion that was described in Kagawa et al. [12] was identified

and its highly repetitive nature [18]. In this study, we showed that a Rotor syndrome-causing LINE-1 insertion in *SLCO1B3* can be easily screened from the investigation of raw sequencing data. In experiments utilizing the SNUH repository, the positive result from LINE-1 chimeric sequence analysis always led to positive experimental results. The agreement of frequency of the *SLCO1B3* LINE-1 insertion in the Japanese population observed by Kagawa et al. [12] and our study validates the correctness of our method.

Although digenic recessive inheritance in Rotor syndrome was established in 2012 [10], only a few studies have reported co-occurring inactivating variants from both *SLCO1B1* and *SLCO1B3* [10, 12, 19]. This is probably because a considerable proportion of the disease-causing variants in these two genes, especially in *SLCO1B3*, are structural variants, such as large deletions, inversions, and insertions of large genetic elements, which cannot be detected by routine sequence analysis [10–12, 19]. This is the first report of genetically diagnosed cases of Rotor syndrome in Korea, and the unique haplotype observed in our study was the same as the only haplotype observed in Japan to date. Although this haplotype, which consists of *SLCO1B1* c.1738C>T, p.R580X, and the *SLCO1B3* intron 6 LINE-1 insertion, is the only haplotype observed in Rotor syndrome patients in Korea and Japan so far, considering the high frequency of *SLCO1B3* intron 6 LINE-1 insertion in East Asia, there is a possibility that other pathogenic variants in *SLCO1B1* co-occur with *SLCO1B3* intron 6 LINE-1 insertion. At least in the East Asian population, *SLCO1B3* intron 6 LINE-1 insertion should be routinely examined during genetic testing for hyperbilirubinemia patients, and the methods proposed in this study would enable this in a simple way.

Zhou et al. [19] reported a Chinese case of Rotor syndrome caused by an exon 4 inversion which was presumably caused by adjacent LINE-1 insertion. Exon 4 in their study is exon 6 when aligned to NM_019844.3 of hg19 and the LINE-1 insertion described in their study shares one junction with the LINE-1 insertion discussed in our study (reverse complemented sequence of 3′ junction of LINE-1 insertion described in Zhou et al. is the same as the 5′ junction sequence in our study). In our study, LINE-1 insertions found from SNUH repository were confirmed either by long-range PCR or gap-PCR tests but the exact configuration of LINE-1 insertions found from 1000 Genomes Project data were not confirmed by additional tests and the positive cases may contain variations of intronic LINE-1 insertion such as the one described in Zhou et al. We additionally evaluated the other junction described in Zhou et al. (hg19 12:21012909, not covered by exome sequencing) from 54 positive cases from 1000 Genomes Project data using low-coverage WGS data and the sign of LINE-1 insertion was not found. However, the evaluation was limited because of the poor coverage (not covered in 32 cases, average depth 4.4 in 22 covered cases). This is one limitation of our study but considering that the LINE-1 insertion described in Zhou et al. also destroys *SLCO1B3* function and causes Rotor syndrome, this would not hamper the importance of evaluation of the intronic LINE-1 insertion location we proposed.

As a transporter protein, *SLCO1B3* has been reported to be associated with the metabolism of various drugs. *SLCO1B3* polymorphisms have been previously linked to the pharmacokinetics of drugs such as imatinib, sulfonylurea, digoxin, and mycophenolic acids [20–23]. In addition, many anti-cancer agents,

Table 2. Frequency estimation of intronic LINE-1 insertion in *SLCO1B3* in Korean population from SNUH repository and read statistics measured from included BAM files

Genotype revealed by confirmatory PCR	No. of patients ($N = 725$)	BAM file findings		
		Read depth	Right-clipped Reads	Estimated VAF of LINE-1 insertion
Heterozygous LINE-1 insertion	91 (12.6%)	218.1	41.6	19.7%
Homozygous LINE-1 insertion	4 (0.6%)	68.9	55.5	81.2%
No insertion	630 (86.9%)	382.6	0.0	0.0%

Estimated allele frequency: 6.8%
BAM binary alignment map, *VAF* variant allele frequency, *LINE-1* long interspersed nuclear elements-1

Table 3. Population-specific frequency of LINE-1 insertion in *SLCO1B3* estimated from the analysis of 1000 Genomes Project data

Super-population	Population	No. of LINE-1 insertion carriers	Total	Carrier frequency (%)	Allele frequency (%)
East Asian Ancestry	Southern Han Chinese	20	108	18.5%	9.3%
East Asian Ancestry	Han Chinese	10	103	9.7%	4.9%
East Asian Ancestry	Japanese	11	104	10.6%	5.3%
East Asian Ancestry	Dai Chinese	7	99	7.1%	3.5%
East Asian Ancestry	Kinh Vietnamese	4	101	4.0%	2.0%
South Asian Ancestry	Tamil	1	103	1.0%	0.5%
South Asian Ancestry	Bengali	1	84	1.2%	0.6%
East Asian Ancestry (Total)		52	515	10.1%	5.0%
South Asian Ancestry (Total)		2	187	1.1%	0.5%

LINE-1 long interspersed nuclear elements-1

including methotrexate [24], paclitaxel [25], docetaxel [26], cisplatin [27], carboplatin [27], irinotecan metabolite SN-38 [28], and certain tyrosine kinase inhibitors [29] are substrates of *SLCO1B3* and there have been recent reports that describe the resistance to anti-cancer drugs caused by abnormal expression or genetic variants of *SLCO1B3* [30]. Specifically, downregulation of *SLCO1B3* was observed in taxane-resistant prostate cancer cell lines [31]. In another study performed by Kagawa et al. [32], hepatic uptake of indocyanine green was severely impaired in individuals homozygous for the same *SLCO1B3* intronic LINE-1 insertion. The impairment of uptake was observed from patients without any *SLCO1B1* variant and this suggests that *SLCO1B3* alone can contribute to the potential effect on drug metabolism. Considering the high frequency of the *SLCO1B3* intron 6 LINE-1 insertion in East Asia, research on the association between this insertion and the response to the drugs mentioned above would be valuable.

Collectively, our findings indicate that the *SLCO1B3* intron 6 LINE-1 insertion is highly frequent and specific in East Asian populations. Considering its frequency and proven pathogenicity, evaluation of this insertion is warranted for suspected Rotor syndrome patients, especially in East Asia.

REFERENCES

1. Ostertag EM, Kazazian HH Jr. Biology of mammalian L1 retrotransposons. Annu Rev Genet. 2001;35:501–38.
2. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409:860–921.
3. Scott AF, Schmeckpeper BJ, Abdelrazik M, Comey CT, O'Hara B, Rossiter JP, et al. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. Genomics. 1987;1:113–25.
4. Szak ST, Pickeral OK, Makalowski W, Boguski MS, Landsman D, Boeke JD. Molecular archeology of L1 insertions in the human genome. Genome Biol. 2002;3:1–18.
5. Sassaman DM, Dombroski BA, Moran JV, Kimberland ML, Naas TP, DeBerardinis RJ, et al. Many human L1 elements are capable of retrotransposition. Nat Genet. 1997;16:37–43.
6. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV, et al. Hot L1s account for the bulk of retrotransposition in the human population. Proc Natl Acad Sci USA. 2003;100:5280–5.
7. Hancks DC, Kazazian HH Jr. Roles for retrotransposon insertions in human disease. Mob DNA 2016;7:9.
8. Zhang X, Zhang R, Yu J. New understanding of the relevant role of LINE-1 retrotransposition in human disease and immune modulation. Front Cell Dev Biol. 2020;8:657.
9. Rotor AB. Familial non-hemolytic jaundice with direct van den Bergh reaction. Acta Med Philip. 1948;5:37–49.
10. van de Steeg E, Stránecký V, Hartmannová H, Nosková L, Hřebíček M, Wagenaar E, et al. Complete OATP1B1 and OATP1B3 deficiency causes human rotor syndrome by interrupting conjugated bilirubin reuptake into the liver. J Clin Investig. 2012;122:519–28.
11. Jirsa M, Knisely A, Schinkel A, Kmoch S. Rotor syndrome. In: GeneReviews® [Internet]. University of Washington, Seattle; 2019.
12. Kagawa T, Oka A, Kobayashi Y, Hiasa Y, Kitamura T, Sakugawa H, et al. Recessive inheritance of population-specific intronic LINE-1 insertion causes a rotor syndrome phenotype. Hum Mutat. 2015;36:327–32.
13. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. Nature. 2015;526:68–74.
14. Stenson PD, Mort M, Ball EV, Evans K, Hayden M, Heywood S, et al. The human gene mutation database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. Hum Genet. 2017;136:665–77.
15. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.
16. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
17. Kim YG, Kim MJ, Lee JS, Lee JA, Song JY, Cho SI, et al. SnackVar: an open-source software for sanger sequencing analysis optimized for clinical use. J Mol Diagn. 2021;23:140–8.
18. Steranka JP, Tang Z, Grivainis M, Huang CRL, Payer LM, Rego FOR, et al. Transposon insertion profiling by sequencing (TIPseq) for mapping LINE-1 insertions in the human genome. Mob DNA. 2019;10:8.
19. Zhou D, Qi S, Zhang W, Wu L, Xu A, Li X, et al. Insertion of LINE-1 retrotransposon inducing exon inversion causes a rotor syndrome phenotype. Front Genet. 2019;10:1399.

20. Yamakawa Y, Hamada A, Nakashima R, Yuki M, Hirayama C, Kawaguchi T, et al. Association of genetic polymorphisms in the influx transporter SLCO1B3 and the efflux transporter ABCB1 with imatinib pharmacokinetics in patients with chronic myeloid leukemia. Ther Drug Monit. 2011;33:244–50.
21. Ren Q, Han X, Ren J, Liu X, Ji L. Influence of the SLCO1B3 gene on sulfonylurea failure in patients with type 2 diabetes in China. Exp Clin Endocrinol Diabetes. 2017;125:449–53.
22. Tounsi N, Trabelsi I, Kerkeni E, Grissa MH, Fredj N, Sekma A, et al. ABCB1 and SLCO1B3 gene polymorphisms and their impact on digoxin pharmacokinetics in atrial fibrillation patients among the tunisian population. Pharmacology. 2017;99:250–8.
23. Tague LK, Byers DE, Hachem R, Kreisel D, Krupnick AS, Kulkarni HS, et al. Impact of SLCO1B3 polymorphisms on clinical outcomes in lung allograft recipients receiving mycophenolic acid. Pharmacogenomics J. 2020;20:69–79.
24. Abe T, Unno M, Onogawa T, Tokui T, Kondo TN, Nakagomi R, et al. LST-2, a human liver-specific organic anion transporter, determines methotrexate sensitivity in gastrointestinal cancers. Gastroenterology. 2001;120:1689–99.
25. Smith NF, Acharya MR, Desai N, Figg WD, Sparreboom A. Identification of OATP1B3 as a high-affinity hepatocellular transporter of paclitaxel. Cancer Biol Ther. 2005;4:815–8.
26. Lee HH, Leake BF, Teft W, Tirona RG, Kim RB, Ho RH. Contribution of hepatic organic anion-transporting polypeptides to docetaxel uptake and clearance. Mol Cancer Ther. 2015;14:994–1003.
27. Lancaster CS, Sprowl JA, Walker AL, Hu S, Gibson AA, Sparreboom A. Modulation of OATP1B-type transporter function alters cellular uptake and disposition of platinum chemotherapeutics. Mol Cancer Ther. 2013;12:1537–44.
28. Yamaguchi H, Kobayashi M, Okada M, Takeuchi T, Unno M, Abe T, et al. Rapid screening of antineoplastic candidates for the human organic anion transporter OATP1B3 substrates using fluorescent probes. Cancer Lett. 2008;260:163–9.
29. Hu S, Franke RM, Filipski KK, Hu C, Orwick SJ, de Bruijn EA, et al. Interaction of imatinib with human organic ion carriers. Clin Cancer Res. 2008;14:3141–8.
30. Sun R, Ying Y, Tang Z, Liu T, Shi F, Li H, et al. The emerging role of the SLCO1B3 protein in cancer resistance. Protein Pept Lett. 2020;27:17–29.
31. de Morrée ES, Böttcher R, van Soest RJ, Aghai A, de Ridder CM, Gibson AA, et al. Loss of SLCO1B3 drives taxane resistance in prostate cancer. Br J Cancer. 2016;115:674–81.
32. Kagawa T, Adachi Y, Hashimoto N, Mitsui H, Ohashi T, Yoneda M, et al. Loss of organic anion transporting polypeptide 1B3 function causes marked delay in indocyanine green clearance without any clinical symptoms. Hepatology. 2017;65:1065–8.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s10038-021-00967-1.

Correspondence and requests for materials should be addressed to M.-W.S.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.