ARTICLE

Application of targeted nanopore sequencing for the screening and determination of structural variants in patients with Lynch syndrome

Kiyoshi Yamaguchi [1] · Rika Kasajima[2,3] · Kiyoko Takane[1] · Seira Hatakeyama[1] · Eigo Shimizu[2] ·
Rui Yamaguchi[2,4,5] · Kotoe Katayama[2] · Masami Arai[6] · Chikashi Ishioka[6] · Takeo Iwama[6] · Satoshi Kaneko [6] ·
Nagahide Matsubara[6] · Yoshihiro Moriya[6] · Tadashi Nomizu[6] · Kokichi Sugano[6] · Kazuo Tamura [6] ·
Naohiro Tomita[6] · Teruhiko Yoshida[6] · Kenichi Sugihara[6] · Yusuke Nakamura[7] · Satoru Miyano[2,8] · Seiya Imoto[2] ·
Yoichi Furukawa [1,6] · Tsuneo Ikenoue[1]

Abstract
Lynch syndrome is a hereditary disease characterized by an increased risk of colorectal and other cancers. Germline variants in the mismatch repair (MMR) genes are responsible for this disease. Previously, we screened the MMR genes in colorectal cancer patients who fulfilled modified Amsterdam II criteria, and multiplex ligation-dependent probe amplification (MPLA) identified 11 structural variants (SVs) of *MLH1* and *MSH2* in 17 patients. In this study, we have tested the efficacy of long read-sequencing coupled with target enrichment for the determination of SVs and their breakpoints. DNA was captured by array probes designed to hybridize with target regions including four MMR genes and then sequenced using MinION, a nanopore sequencing platform. Approximately, 1000-fold coverage was obtained in the target regions compared with other regions. Application of this system to four test cases among the 17 patients correctly mapped the breakpoints. In addition, we newly found a deletion across an 84 kb region of *MSH2* in a case without the pathogenic single nucleotide variants. These data suggest that long read-sequencing combined with hybridization-based enrichment is an efficient method to identify both SVs and their breakpoints. This strategy might replace MLPA for the screening of SVs in hereditary diseases.

Introduction

Lynch syndrome (LS, [MIM: 120435]), previously referred to as hereditary non-polyposis colorectal cancer (HNPCC), is the most common form of hereditary colorectal cancer (CRC), accounting for 2–4% of new CRC diagnoses [1]. Apart from CRCs, affected individuals have an increased

These authors contributed equally: Kiyoshi Yamaguchi, Rika Kasajima

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s10038-021-00927-9.

✉ Yoichi Furukawa
   yofurukawa@g.ecc.u-tokyo.ac.jp

1   Division of Clinical Genome Research, Advanced Clinical Research Center, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan

2   Division of Health Medical Intelligence, Human Genome Center, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan

3   Molecular Pathology and Genetics Division, Kanagawa Cancer Center Research Institute, Kanagawa 241-8518, Japan

4   Division of Cancer Systems Biology, Aichi Cancer Center Research Institute, Aichi 464-8681, Japan

5   Division of Cancer Informatics, Nagoya University Graduate School of Medicine, Aichi 466-8550, Japan

6   The Committee of HNPCC Registry and Genetic Testing Project, Japanese Society for Cancer of the Colon and Rectum (JSCCR), Tokyo 102-0075, Japan

7   Cancer Precision Medicine Center, Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan

8   Systems Biology for Intractable Diseases, Medical Research Institute, Tokyo Medical and Dental University, Tokyo 113-8510, Japan

risk of developing other types of cancer such as endometrial, renal pelvic and ureteral, gastric, and small intestinal cancers. LS is caused by germline variants in mismatch repair (MMR) genes including mutL homolog 1 (*MLH1*), mutS homolog 2 (*MSH2*), mutS homolog 6 (*MSH6*), and PMS1 homolog 2, MMR system component (*PMS2*). In addition, deletions within the 3' end of the *EPCAM* gene have been shown to cause LS through the inactivation of the *MSH2* promoter [2].

LS is mainly diagnosed by three approaches: (1) testing the tumors for microsatellite instability or MMR proteins by immunohistochemistry to identify patients who should undergo germline testing, (2) genetic testing for germline variants in patients who have a family history or have suspected LS, or (3) as a secondary finding in cancer profile analysis for precision medicine [3]. In the case of genetic testing, variants in the major MMR genes have been analyzed by a panel sequencing to detect small alterations such as single nucleotide variants (SNVs) and short insertions/deletions (indels), and copy number variations (CNVs). Multiplex ligation-dependent probe amplification (MLPA) has been also applied for the detection of large deletions/duplications. However, MLPA does not determine the breakpoints of SVs. For the cascade screening of family members of a proband with a pathogenic SV, the identification of breakpoint is useful because polymerase chain reaction (PCR) amplification or direct sequencing can be used for the detection of SVs. In addition, inversions and SVs in deep intronic regions are not detected by MLPA.

As a corroborative project with the Japanese society for cancer of the colon and rectum "HNPCC registry and genetic testing project", we previously analyzed *MLH1*, *MSH2*, and *MSH6* by direct sequencing and MLPA in 111 Japanese patients with CRC who fulfilled the modified Amsterdam II criteria including gastric cancer as LS-related cancer [4]. This study identified a total of 55 pathogenic variants in 62 patients. It is of note that among the 55 variants, 11 were structural variants (SVs) including 6 exonic deletions and 5 duplications. Although we determined the breakpoints for eight of the 11 SVs by long-range-PCR followed by Sanger sequencing, we were unable to map the breakpoints in the remaining three cases.

Since next-generation sequencing (NGS) dramatically improved the output of sequencing capacity, genetic diagnosis using gene panels has been replacing genetic tests using the Sanger sequencing. NGS-based multigene panels are available for the screening of germline variants in the *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM* genes [5]. Multigene panel sequencing or whole-exome sequencing can lead to the isolation of large deletions or duplications, but these methodologies cannot identify the breakpoints, and they miss complex SVs and CNV-neutral rearrangements such as inversions and large intronic insertions [6].

Therefore, panel or whole-exome sequencing alone might not be enough to provide a conclusive result for genetic diagnosis. Ideally, it is desirable to identify both small nucleotide changes and breakpoints of SVs in the same system. Since the nanopore-based sequencing technology has enabled long-read sequencing, we applied MinION, a nanopore DNA sequencer developed by Oxford Nanopore Technologies (Oxford, UK), in combination with a hybridization-based target enrichment (SureSelect, Agilent Technologies, Santa Clara, CA). We first assessed the efficacy of target enrichment using four LS cases carrying SVs whose breakpoints had already been characterized in a previous study and then tested a case in which we did not find any pathogenic variants in the MMR genes by PCR-direct sequencing. Our data demonstrate that long read-sequencing combined with hybridization-based enrichment is a useful method for the identification of deletion/duplication breakpoints.

Materials and methods

Ethical approval and DNA samples

This project was approved by the ethical committee of the Institute of Medical Science, the University of Tokyo (IMSUT-IRB:28-27). Genomic DNAs used in a previous collaborative study with the Japanese Society for Cancer of the Colon and Rectum [4] were analyzed using the MinION sequencing. Written informed consent was obtained from all the patients participating in the project.

Capture probes for the enrichment of genomic regions of the *MMR* and *EPCAM* genes

We designed 120-mer RNA capture probes for the entire genomic sequence (exons including 5'- and 3'-untranslated regions and introns) of the *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM* genes, and the intergenic region between the *EPCAM* and *MSH2* genes using Agilent SureDesign (Agilent Technologies, Santa Clara, CA, https://earray.chem.agilent.com/suredesign/). The probes designed (9,785 probes) were estimated to cover 58.9% of the targeted region on average (Fig. 1a).

MinION sequencing and data analysis

Sequencing libraries were prepared using Ligation Sequencing Kit 1D (SQK-LSK109, Oxford Nanopore Technologies, Oxford, UK). Briefly, 3 μg of DNA were digested with dsDNA fragmentase (NEBNext dsDNA Fragmentase, New England Biolabs, Ipswich, MA). After PCR amplification using LongAmp Taq DNA Polymerase

Fig. 1 Enrichment of the MMR genes using RNA capture probes. **a** The regions expected to be covered by capture probes (green) and the coverage achieved using nanopore sequencing of JLS058 (emerald green). **b** Locations of four SVs including three deletions (red) and a duplication (green) in the *MLH1* and *MSH2* genes. **c** Proportion of the MinION sequencing reads mapped to reference genome (hg19). Reads were classified into nine groups according to read length. Data are represented as the mean ± SD. **d** Proportion of reads correctly mapped to the targeted regions (histogram) and coverage of the regions by the reads (line graph). The percentage of on-target reads and mean depth of the target regions was separately calculated for each group. Data are represented as the mean ± SD. **e** Snapshots of the Integrative Genomics Viewer depicting the four pathogenic SVs
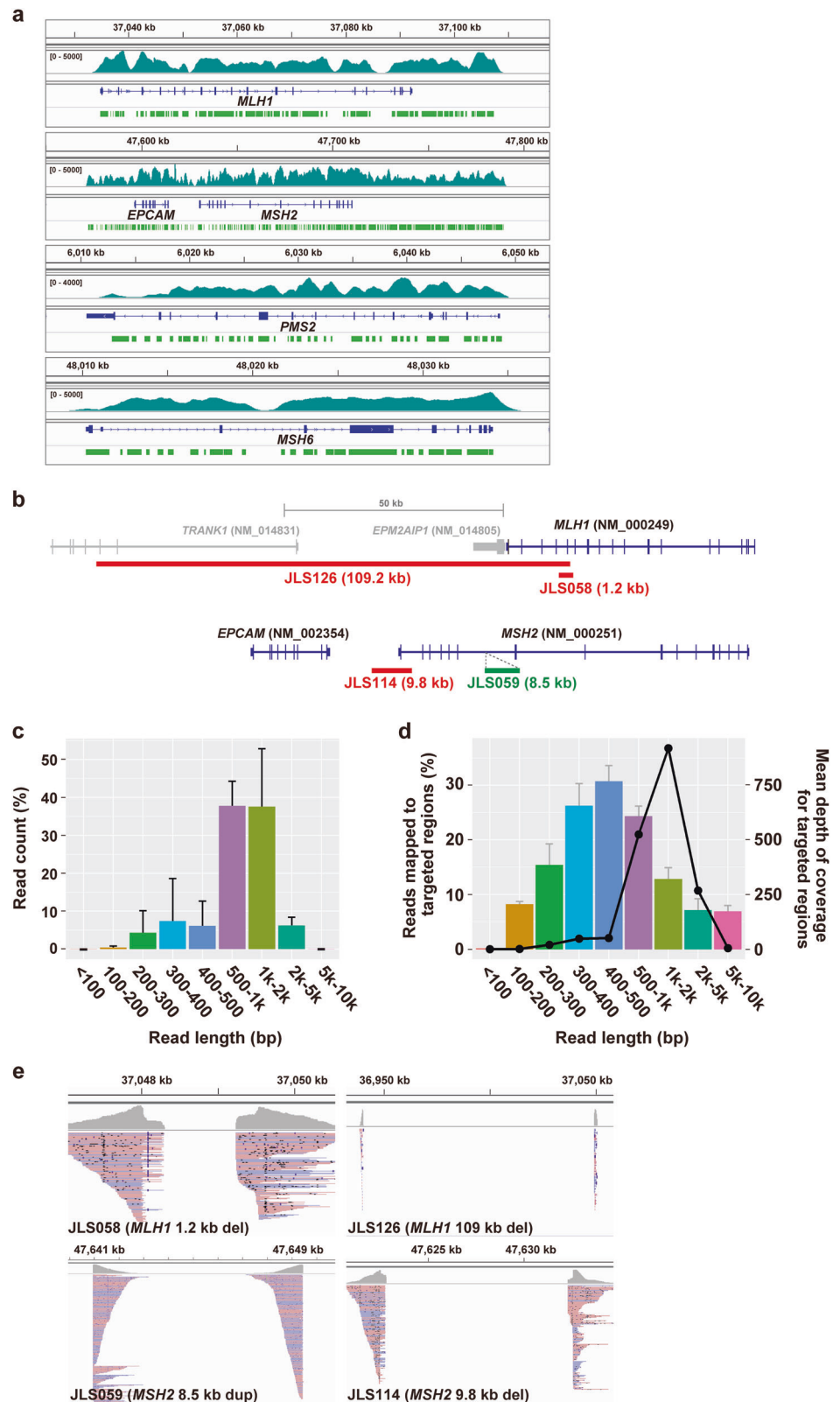
(New England Biolabs), genomic regions of the *MMR* and *EPCAM* genes were captured using the SureSelect Custom DNA Target Enrichment Probes (Agilent Technologies).

Subsequently, sequencing libraries were generated by additional PCR amplification and adapter ligation. Each library was loaded onto a R9.4 flow cell (Oxford Nanopore

Technologies) and sequenced for 12–24 h. The basecalling of the raw data was performed using Guppy (ver.2.3.5 + 53a111f, Oxford Nanopore Technologies). Quality was assessed using MinIONQC [7] (ver.1.4.1). Long-read alignment and SV detection were performed using NGMLR-align algorithm (ver.0.2.7) [8] and Sniffles (ver.1.0.8) [8], respectively. Sequencing reads mapped outside the target regions were removed. The candidate variants were further filtered by removing variants with both VAF < 0.1 and reads <20. In addition, Minimap2 (ver.2.12) [9] and NanoVar (ver.1.3.2) [10] tools were also used for read alignment and SV detection, respectively.

PCR-direct sequencing and MLPA

DNA fragments including breakpoints were amplified by PCR using KOD One (TOYOBO, Osaka, Japan), and the PCR products were sequenced on the Applied Biosystems 3500xl DNA Analyzer using the BigDye Direct Cycle Sequencing kit (Thermo Fisher Scientific, Waltham, MA). The primer sequences used for the amplification and sequencing are shown in Supplementary Table S1. MLPA was performed using a Salsa MLPA kit (MRC Holland, Amsterdam, Netherlands) in FALCO biosystems (Kyoto, Japan).

Results

Enrichment of target DNA by a hybridization-based capture

Among the 11 LS cases with SVs, we chose four cases carrying a wide range of deletions and a duplication; three cases with 1.2 kb (JLS058) and 109.2 kb (JLS126) deletions in *MLH1*, and a 9.8 kb (JLS114) deletion in *MSH2*, and a case with an 8.5 kb (JLS059) duplication in *MSH2* (Fig. 1b).

Initially, we applied the MinION sequencing for JLS126 without a targeted enrichment strategy. Although the MinION sequencing produced 6.6 Gb of data (mean read length: 5614 bp and max. read length: 184,450 bp), this analysis failed to detect the deletion in *MLH1* (c.1-94968_c.453 + 696del109180) due to the low coverage of the deleted region (average depth of 2.1× by estimation). To increase the coverage, we combined a targeted enrichment with the MinION sequencing and designed a total of 9,785 RNA probes scattered in the entire genomic regions (exons, introns, and intergenic regions) of the *MLH1*, *MSH2*, *MSH6*, *PMS2*, and *EPCAM* genes (Fig. 1a). The coverage of the probes to the targeted regions was 58.9%. Although it is generally recommended to shear DNA into 150–200 bp in size for the hybridization to SureSelect probes, we digested DNA into 1–3 kb using dsDNA fragmentase to assess the

enrichment efficiency of DNA fragments of this length. The sequencing of four cases (JLS058, JLS126, JLS114, and JLS059) achieved an average throughput of 6.5 Gb (3.4–9.4 Gb), and the average length of read was 1228 bp (913–1421 bp) (Supplementary Table S2). For JLS126, the total output was 3.4 Gb that corresponds to approximately 1.1× haplotype genome, but the average depth of the target regions was as high as 1201× by the enrichment, indicating 1099-fold enrichment of the target regions. The enrichment observed in JLS058, JLS114, and JLS059 was 853-, 767- and 962-fold, respectively. These data suggested that the capture system could enrich the target regions by approximately 1000-fold compared with the other regions.

We further analyzed the efficiency of enrichment by the read length in the four cases. The percentage of the total reads for read lengths 300–400 bp, 400–500 bp, 500–1000 bp, and 1–2 kb were 7.4%, 6.1%, 37.8%, and 37.6%, respectively, and those shorter than 300 bp and longer than 2 kb were 4.8% and 6.3%, respectively (Fig. 1c). Interestingly, the ratio of reads mapped to the target regions was different according to the length of reads. The ratios for read lengths 300–400 bp, 400–500 bp, 500–1000 bp, and 1–2 kb were 26.2%, 30.7%, 24.3%, and 12.8%, respectively (Fig. 1d). Although reads with a length of 1–2 kb were less aligned than those of 500 bp–1 kb, the former accounted for more coverage of the target regions than the latter (Fig. 1d, line chart). These data suggest that the RNA probes can enrich DNA fragments up to 2 kb, and that captured DNA with length between 500 bp and 2 kb efficiently cover the target regions.

Detection and accurate mapping of SVs using nanopore sequencing

To isolate the SVs and their breakpoints in the four test cases, we used NGMLR-Sniffles because the combination of aligner and structural variation caller has been reported to provide accurate results in SV analysis [8]. Sniffles considers the consistency of the breakpoint positions, and reports the average positions that is classified as either "Precise" or "Imprecise". As a result, this pipeline with an additional filtration detected single SVs in JLS058 and JLS114, whereas it called seven and twelve SVs in JLS126 and JLS059, respectively. As expected, the list of SVs included the four pathogenic SVs (Table 1). The pathogenic variants in JLS126 and JLS059 were reported as "Precise", and their breakpoints were correctly mapped. On the other hand, the breakpoint positions of JLS058 and JLS114 were reported as "Imprecise", and shifted by 29–258 bp (Table 1). The information of either "Precise" or "Imprecise" structural variants likely depends on the sequences surrounding the breakpoints. "Precise" variants correspond to SVs whose 5′- or 3′-breakpoint positions were partly

Table 1 List of SVs called by NGMLR-Sniffles.

Sample	CHR1	POS	CHR2	END	SVLEN	ALT	Precise/imprecise	DR	DV	AF	VAL POS	VAL LEN
JLS058	**chr3**	**37,048,022**	**chr3**	**37,049,242**	**1220**	**DEL**	**IMPRECISE**	**612**	**135**	**0.18**	**chr3: 37,048,051-37,049,271**	**1,221**
JLS126	chr2	47,689,038	chr2	47,689,080	42	DEL	PRECISE	11	22	0.67		
	chr2	47,734,176	chr2	47,734,221	45	DEL	IMPRECISE	168	29	0.15		
	chr2	47,617,606	chr2	70,800,134	–	TRA	PRECISE	4	47	0.92		
	chr2	47,617,607	chr2	117,086,264	–	TRA	PRECISE	5	35	0.88		
	chr2	47,617,813	chr2	34,329,765	–	TRA	PRECISE	196	29	0.13		
	chr2	47,617,826	chr2	95,340,743	–	TRA	PRECISE	247	28	0.10		
	chr3	**36,940,066**	**chr3**	**37,049,244**	**109,178**	**DEL**	**PRECISE**	**121**	**28**	**0.19**	**chr3: 36,940,071-37,049,250**	**109,180**
JLS059	chr2	47,610,746	chr2	47,610,777	31	DEL	PRECISE	275	95	0.26		
	chr2	**47,649,462**	**chr2**	**47,640,953**	**8509**	**DUP**	**PRECISE**	**627**	**246**	**0.28**	**chr2: 47,640,953-47,649,462**	**8,510**
	chr2	47,648,219	chr2	47,648,275	56	DEL	IMPRECISE	352	58	0.14		
	chr2	47,654,110	chr2	47,654,192	82	DEL	IMPRECISE	571	321	0.36		
	chr2	47,654,092	chr2	47,654,168	76	DEL	IMPRECISE	571	274	0.32		
	chr2	47,662,644	chr2	47,662,682	38	DEL	PRECISE	172	25	0.13		
	chr2	47,688,876	chr2	47,688,908	32	DEL	IMPRECISE	883	186	0.17		
	chr2	47,708,915	chr2	47,708,961	46	DEL	IMPRECISE	276	36	0.12		
	chr2	47,734,125	chr2	47,734,231	106	DEL	IMPRECISE	411	58	0.12		
	chr2	48,016,026	chr2	48,016,078	52	DEL	PRECISE	172	114	0.40		
	chr3	37,039,537	chr3	37,039,576	39	DEL	PRECISE	18	77	0.81		
	chr7	6,033,482	chr7	6,033,519	37	DEL	PRECISE	97	21	0.18		
JLS114	**chr2**	**47,622,523**	**chr2**	**47,632,305**	**9782**	**DEL**	**IMPRECISE**	**596**	**128**	**0.18**	**chr2: 47,622,781-47,632,560**	**9780**

Bold letter represents the pathogenic variants.

CHR1 chromosome name where SV occurred, *POS* starting breakpoint position of SV, *CHR2* chromosome of second breakpoint of SV, *END* position of the second breakpoint of SV, *SVLEN* length of SV, *ALT* types of SV, *DEL* deletion, *DUP* duplication, *TRA* translocation, *IMPRECISE/PRECISE* confidence of the exact breakpoint position, *DR* reference read, *DV* variant read, *AF* allele frequency, *VAL POS* validated position of the breakpoint, *VAL LEN* validated length of SV

outside of interspersed repeats and low complexity DNA sequences. On the other hand, both 5′- and 3′-breakpoint positions of "Imprecise" variants are located within the repeats or low complexity DNA sequences (data not shown, RepeatMasker, http://www.repeatmasker.org/). Although we tested Minimap2-NanoVar [9, 10], another combination of aligner and SV caller, these algorithms failed to call the four pathogenic SVs. Since Minimap2-Sniffles also detected these pathogenic SVs, Sniffles might be appropriate for calling SVs in the data obtained by this experimental condition. In the analysis of nanopore sequencing reads by NGMLR-Sniffles, we found that the breakpoint position was determined with the accuracy of 1.0 ± 2.0 bp in "Precise" and 142.8 ± 131.4 bp in "Imprecise", and the size was determined with that of 2.0 ± 1.4 bp (Fig. 1e). These data are useful for the design of primer set to amplify the variant DNA fragments and subsequent validation of the exact breakpoints. It is of note that the variant frequency of reads carrying the SVs were 0.19, 0.28, 0.18, and 0.18 in

JLS126, JLS059, JLS114, and JLS058, respectively, suggesting that reads containing the SVs were less efficiently mapped to the position, compared with those without SVs. These low-frequencies of SVs might be due to the locations of capture probes. Because introns and intergenic regions often harbor repetitive sequences such as LINEs and SINEs, probes are usually designed outside of such sequences for sequence specificity. On the other hand, many SVs are associated with repetitive sequences because they are generated during DNA recombination or repair [11, 12]. Therefore, an efficient capture system for longer DNA fragments is essential to increase efficacy.

We further analyzed 13 other deletions (2 in JLS126 and 11 in JLS059) and found that all were mid-range in size (31–106 bp) with highly repetitive sequence, implying characteristics of error-prone reads by nanopore sequencing, or just PCR error during the amplification of DNA libraries. Notably, these deletions will not be judged as pathogenic SVs because they are located within deep intronic

regions of the MMR genes or the 3'-flanking region of *MSH2*, and are assumed not to disrupt the transcripts. In the future, accumulation of data or improvement of analytical pipelines might enable us to correctly evaluate these variants.

Screening of structural variants in a case without pathogenic SNVs in the MMR genes

To evaluate the efficacy of MinION sequencing combined with the target enrichment, we analyzed JLS128, a case enrolled in the collaborative project. In this sample, we previously identified a three-base deletion c.279_281delTCT in *MSH2* [4], which had been reported to be associated with splicing defects [13]. However, the interpretation has been changed and this variant is now judged as benign in ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/). Therefore, we screened for SVs using the targeted MinION sequencing. As shown in Fig. 2a, analysis of the sequencing data detected a deletion in *MSH2* and mapped the breakpoint of the deletion spanning 84,126 bp in chr2: 47,651,634–47,735,760 (GRCh37). Subsequent PCR-based Sanger sequencing confirmed the size (84,128 bp) and location (chr2: 47,651,635–47,735,762) of the breakpoints with a slight positional shift (NM_000251.3(*MSH2*):c.1077-5246_*+256 74del, Fig. 2b). In agreement with this result, MLPA detected a large deletion encompassing exons 7–16 of *MSH2* (Fig. 2c).

Discussion

In LS, it has been reported that large deletions accounted for 28.6% of pathogenic variants in the *MLH1* and *MSH2* genes in the United States [14]. Large deletions significantly contributed and accounted for 20–22% of the mutations in *MLH1*, *MSH2*, and *PMS2*, and 4% in *MSH6* in Swedish LS [15]. Of note, genomic rearrangements are likely prevalent events in *MSH2*, where they comprise up to 45.3% of mutations [16], indicating that screening of structural alterations in the MMR genes is necessary for the genetic diagnosis of this disease. For the identification of structural alterations, Southern blotting [16, 17], long-range PCR [18], quantitative PCR [19], and MLPA [16, 18] have been employed. In this report, we have shown that targeted MinION sequencing for multiple genes is useful for the identification of the breakpoints as well as pathogenic large SVs. Targeted capture nanopore sequencing has also been applied for the detection of somatic structural variants in *ABCB1*, and determined translocations in the promoter region of *ABCB1* in AML cells [20]. Although our study successfully identified the breakpoints of five pathogenic SVs, there remains a challenge of false positives detected by current algorithms including NGMLR-Sniffles. Since probes are not designed within repetitive sequences, sequence reads containing breakpoint in highly repetitive sequences may not be captured by hybridization-based enrichment. In addition, probes located far from the target
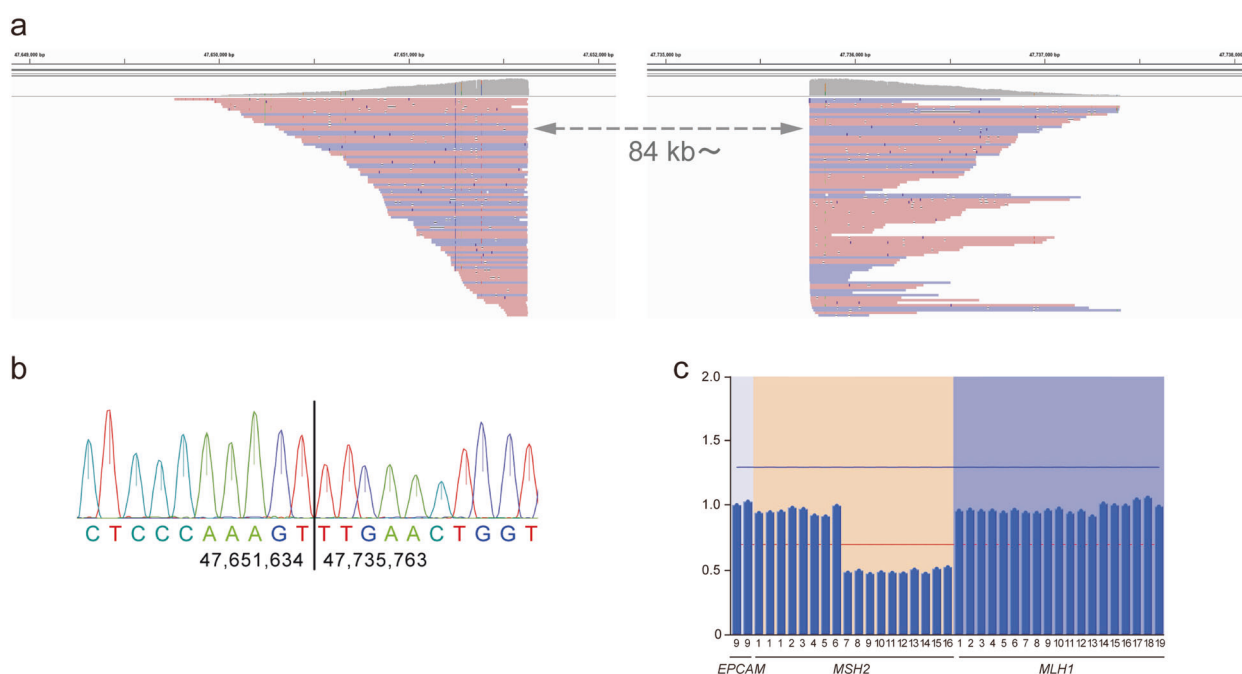
Fig. 2 Identification of *MSH2* deletion in JLS128. **a** A snapshot of the Integrative Genomics Viewer depicting a deletion of *MSH2* in JLS128. Mutant reads mapped to the forward and reverse strands are shown in red and blue, respectively. Blue and black bars indicate indels (small indels less than 3 bp were filtered out). **b** Eighty-four kilobase deletion spanning exon 7–16 of *MSH2* was confirmed by PCR-direct sequencing. **c** MLPA analysis corroborated the deletion in *MSH2*

regions are not usually synthesized, it is possible to overlook SVs whose breakpoints are all outside of the target regions. Therefore, efficient target enrichment other than coding sequences remains challenging. Recently, a new method termed nanopore Cas9-targeted sequencing was established for the enrichment of target regions [21]. This strategy uses targeted cleavage with specific guide RNA and Cas9 for ligating nanopore sequencing adapters, and enable the capture of more than 10 kb DNA fragments. Furthermore, using Oxford Nanopore Technologies, a method to sequence-specific molecules without prior enrichment is being developed. This approach is called Read Until that allows us to computationally captutre target regions [22, 23]. These methods might enhance the sensitivity and specificity of SV detection.

Analysis of variations in *PMS2* is complex due to the presence of multiple pseudogenes and the gene conversion events with *PMS2CL*, which leads to underestimation of *PMS2* variant carriers [24–26]. Analysis of larger genome has significant advantages over traditional short-read sequencing technology. In addition to long-read sequencing, nanopore technology allows the detection of epigenetic modifications. Since constitutional epimutations of responsible genes are also a possible mechanism for hereditary diseases [27], detection of epigenetic alterations in the associated genes should increase the sensitivity of genetic diagnosis. Nanopore technologies directly read DNA modification without chemical pretreatment such as sodium bisulfite. If the PCR-free method is incorporated in the preparation of the DNA library, Nanopore sequencing will also contribute to the detection of causative epigenetic changes for a wide range of diseases.

Conflict of interest The authors declare no competing interests.

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

References

1. Barrow E, Hill J, Evans DG. Cancer risk in Lynch Syndrome. Fam Cancer. 2013;12:229–40.
2. Ligtenberg MJ, Kuiper RP, Chan TL, Goossens M, Hebeda KM, Voorendt M, et al. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3′ exons of TACSTD1. Nat Genet. 2009;41:112–7.
3. Yurgelun MB, Hampel H. Recent advances in Lynch syndrome: diagnosis, treatment, and cancer prevention. Am Soc Clin Oncol Educ Book. 2018;38:101–09.
4. Ikenoue T, Arai M, Ishioka C, Iwama T, Kaneko S, Matsubara N, et al. Importance of gastric cancer for the diagnosis and surveillance of Japanese Lynch syndrome patients. J Hum Genet. 2019;64:1187–94.
5. Pritchard CC, Smith C, Salipante SJ, Lee MK, Thornton AM, Nord AS, et al. ColoSeq provides comprehensive lynch and polyposis syndrome mutational analysis using massively parallel sequencing. J Mol Diagn. 2012;14:357–66.
6. Morak M, Steinke-Lange V, Massdorf T, Benet-Pages A, Locher M, Laner A, et al. Prevalence of CNV-neutral structural genomic rearrangements in MLH1, MSH2, and PMS2 not detectable in routine NGS diagnostics. Fam Cancer. 2020;19:161–67.
7. Lanfear R, Schalamun M, Kainer D, Wang W, Schwessinger B. MinIONQC: fast and simple quality control for MinION sequencing data. Bioinformatics 2019;35:523–25.
8. Sedlazeck FJ, Rescheneder P, Smolka M, Fang H, Nattestad M, von Haeseler A, et al. Accurate detection of complex structural variations using single-molecule sequencing. Nat Methods. 2018;15:461–68.
9. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 2018;34:3094–100.
10. Tham CY, Tirado-Magallanes R, Goh Y, Fullwood MJ, Koh BTH, Wang W, et al. NanoVar: accurate characterization of patients' genomic structural variants using low-depth nanopore sequencing. Genome Biol. 2020;21:56.
11. Li L, McVety S, Younan R, Liang P, Du Sart D, Gordon PH, et al. Distinct patterns of germ-line deletions in MLH1 and MSH2: the implication of Alu repetitive element in the genetic etiology of Lynch syndrome (HNPCC). Hum Mutat. 2006;27:388.
12. van der Klift H, Wijnen J, Wagner A, Verkuilen P, Tops C, Otway R, et al. Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible for hereditary nonpolyposis colorectal cancer (HNPCC). Genes Chromosomes Cancer. 2005;44:123–38.
13. Tournier I, Vezain M, Martins A, Charbonnier F, Baert-Desurmont S, Olschwang S, et al. A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Hum Mutat. 2008;29:1412–24.
14. Wagner A, Barrows A, Wijnen JT, van der Klift H, Franken PF, Verkuijlen P, et al. Molecular analysis of hereditary nonpolyposis colorectal cancer in the United States: high mutation detection rate among clinically selected families and characterization of an American founder genomic deletion of the MSH2 gene. Am J Hum Genet. 2003;72:1088–100.
15. Lagerstedt-Robinson K, Rohlin A, Aravidis C, Melin B, Nordling M, Stenmark-Askmalm M, et al. Mismatch repair gene mutation spectrum in the Swedish Lynch syndrome population. Oncol Rep. 2016;36:2823–35.
16. Baudhuin LM, Ferber MJ, Winters JL, Steenblock KJ, Swanson RL, French AJ, et al. Characterization of hMLH1 and hMSH2 gene dosage alterations in Lynch syndrome patients. Gastroenterology 2005;129:846–54.
17. Wijnen J, van der Klift H, Vasen H, Khan PM, Menko F, Tops C, et al. MSH2 genomic deletions are a frequent cause of HNPCC. Nat Genet. 1998;20:326–8.
18. Nakagawa H, Hampel H, de la Chapelle A. Identification and characterization of genomic rearrangements of MSH2 and MLH1 in Lynch syndrome (HNPCC) by novel techniques. Hum Mutat. 2003;22:258.
19. Vaughn CP, Lyon E, Samowitz WS. Confirmation of single exon deletions in MLH1 and MSH2 using quantitative polymerase chain reaction. J Mol Diagn. 2008;10:355–60.

20. Williams MS, Basma NJ, Amaral FMR, Williams G, Weightman JP, Breitwieser W, et al. Targeted nanopore sequencing for the identification of ABCB1 promoter translocations in cancer. BMC Cancer. 2020;20:1075.

21. Gilpatrick T, Lee I, Graham JE, Raimondeau E, Bowen R, Heron A, et al. Targeted nanopore sequencing with Cas9-guided adapter ligation. Nat Biotechnol. 2020;38:433–38.

22. Loose M, Malla S, Stout M. Real-time selective sequencing using nanopore technology. Nat Methods. 2016;13:751–4.

23. Payne A, Holmes N, Clarke T, Munro R, Debebe BJ, Loose M. Readfish enables targeted nanopore sequencing of gigabase-sized genomes. Nat Biotechnol. 2021;39:442–50.

24. Hendriks YM, Jagmohan-Changur S, van der Klift HM, Morreau H, van Puijenbroek M, Tops C, et al. Heterozygous mutations in PMS2 cause hereditary nonpolyposis colorectal carcinoma (Lynch syndrome). Gastroenterology 2006;130:312–22.

25. Jansen AML, Tops CMJ, Ruano D, van Eijk R, Wijnen JT, Ten Broeke S, et al. The complexity of screening PMS2 in DNA isolated from formalin-fixed paraffin-embedded material. Eur J Hum Genet. 2020;28:333–38.

26. van der Klift HM, Tops CM, Bik EC, Boogaard MW, Borgstein AM, Hansson KB, et al. Quantification of sequence exchange events between PMS2 and PMS2CL provides a basis for improved mutation scanning of Lynch syndrome patients. Hum Mutat. 2010;31:578–87.

27. Hitchins MP. Constitutional epimutation as a mechanism for cancer causality and heritability? Nat Rev Cancer. 2015;15:625–34.