Received: 1 June 2020 | Revised: 13 August 2020 | Accepted: 30 August 2020

DOI: 10.1002/mgg3.1523

Molecular Genetics & Genomic Medicine Open Access WILEY

ORIGINAL ARTICLE

Insertion of an Alu-like element in *MLH1* intron 7 as a novel cause of Lynch syndrome

Yirong Li[1] | **Erin Salo-Mullen**[2] | **Anna Varghese**[2] | **Magan Trottier**[2] | **Zsofia K. Stadler**[2] | **Liying Zhang**[1,3]

[1]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[2]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[3]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California at Los Angeles (UCLA), Los Angeles, CA, USA

Correspondence
Liying Zhang, Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California at Los Angeles (UCLA), 10833 Le Conte Ave, Los Angeles, CA 90095, USA.
Email: LiyingZhang@mednet.ucla.edu

Funding information
This study was funded by Department of Pathology, Memorial Sloan Kettering Cancer Center.

Abstract

Background: Lynch Syndrome (LS) is caused by germline mutations in the DNA mismatch repair (MMR) genes with mutations in *MLH1* accounting for ~40% of LS-related alterations.

Methods: MSK-IMPACT analysis was performed on peripheral blood from a patient with early- onset colorectal cancer. Subsequently PCR and sequencing was performed to characterize the insertion. Immunohistochemistry for MMR genes and *MLH1* promoter methylation were analyzed on patient's tumor.

Results: MSK-IMPACT germline testing revealed an insertion into c.588+8_588+9 of *MLH1* intron 7. The insertion was further characterized as an AluSx-like element with ~115 bp in length. Functional studies demonstrated that the AluSx-like element led to complete disruption of mRNA splicing and probably resulted in transcriptional termination at the poly (A) region of the AluSx-like insertion.

Conclusions: The insertion of a truncated AluSx like element into *MLH1* intron 7 results in aberrant splicing and transcription, thereby causing Lynch syndrome. This study confirms that retrotransposon insertions may be an important mechanism for cancer predisposition.

KEYWORDS

AluSx, Lynch syndrome, *MLH1*, splicing

1 | INTRODUCTION

Lynch syndrome (LS) is an autosomal dominant cancer predisposition syndrome characterized by an increased risk of many types of cancer, especially colorectal and endometrial cancers (Seth et al., 2018). LS is caused by germline mutations in mismatch repair (MMR) genes. *MLH1* (OMIM # 120436) is one of the most frequently altered genes in Lynch syndrome patients' accountings for 40%–50% of all LS patients (Bonadona et al., 2011; Truninger et al., 2005).

Retrotransposons are genetic transposable elements that can be integrated elsewhere in a genome through an RNA intermediate mechanism. There are two subclasses of retrotransposons, the long terminal repeat (LTR-retrotransposons) and the non-LTR retrotransposons. The non-LTR retrotransposons include long interspersed elements (LINEs) and short interspersed elements (SINEs) which account for approximately one-third of the human genome (Cordaux & Batzer, 2009; Dewannieux et al., 2003). Alu elements are the most abundant short

interspersed elements (SINEs). Alu insertions into protein-coding region have been reported as pathogenic variants in many genes, such as *APC*, *ATM*, *BRCA2*, *BRCA1*, and *CHEK2* (Qian et al., 2017). The Alu insertion within exon 3 of *BRCA2* (c.156_157insAlu), resulting in exon 3 skipping, has been established to contribute to Hereditary Breast and Ovarian Cancer (HBOC) syndrome (Teugels et al., 2005). Recently, Alu element insertion in the *MLH1* exon 6 coding sequence (c.512_513insAlu) was reported as a mutation predisposing to Lynch syndrome (Solassol et al., 2019). In addition to the disruption of protein-coding region, Alu elements can also affect mRNA splicing and result in disease-causing alleles. An Alu insertion into the canonical splicing site of BRCA2 c.8953+1_8953+2 has been reported in a few patients who underwent clinical hereditary cancer genetic testing (Qian et al., 2017). Insertion of a SVA (SINE–VNTR–Alu) element into PMS2 intron 7 c.804-60_804-59 has been reported to alter mRNA splicing as a cause of Lynch syndrome (van der Klift et al., 2012). Most recently, a child with a rare and fatal neurodegenerative genetic disease, neuronal ceroid lipofuscinosis 7 (CLN7, a form of Batten's disease), caused by the insertion of SVA into the intron 6 of MFSD8 was treated clinically with patient-customized oligonucleotide therapy (Kim et al., 2019).

Here, we report that insertion of an AluSx-like element into c.588+8_588+9 of *MLH1* intron 7 caused abnormal mRNA splicing of *MLH1* exon 7, introducing the AluSx-like sequence into the mRNA and resulting in transcriptional termination at the poly (A) region of the AluSx-like insertion as a novel cause of Lynch syndrome.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

The patient in this study consented to an Institutional Review Board (IRB)-approved protocol at Memorial Sloan Kettering Cancer Center.

2.2 | Subject

A 42-year-old Chinese man was diagnosed with metastatic ascending colon moderate differentiated invasive adenocarcinoma. Both parents are alive with no known cancer diagnosis. There is a paternal uncle who died in his 40 s with a reported lung cancer (Figure 1a). The patient was seen at Memorial Sloan Kettering Cancer Center and MSK-IMPACT (IMPACT stands for Integrated Mutation Profiling of Actionable Cancer Targets) testing was performed on his tumor and matched normal specimens.

2.3 | NGS and PCR analysis

The patient tumor and matched normal specimens were tested using the Memorial Sloan Kettering Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT) Germline and somatic sequencing and variant calling were performed as described previously (Cheng et al., 2015, 2017). Germline variants were reviewed by board-certified molecular pathologist or clinical molecular genetics and classified based on the American College of Medical Genetics (ACMG) criteria (Richards et al., 2015). Microsatellite instability was analyzed as previously reported (Middha et al., 2017). PCR for *MLH1* (NM_000249.3) was performed with forward primer F1 5′-GTAAAACGACGGCCAGTAACTA AAAGGGGGCTCTGAC-3′ and reverse primer R1 5′-CAGG AAACAGCTATGACAAATAATGTGATGGAATGATA AACC-3′ under the condition of denature at 96°C for 5 min, 35 cycles (94°C for 30 s, 58°C for 45 s, 72°C for 60 s), and extension at 72°C for 5 min.

3 | RESULTS

MSK-IMPACT testing on patient's colorectal cancer tissue indicated microsatellite instability (MSI) unstable (MSI-H) with MSI sensor score of 26.2 (Middha et al., 2017). IHC showed loss of PMS2 and heterogenous positive staining for MLH1. *MLH1* promoter methylation, which may lead to MLH1/PMS2 protein loss, was not detected in the tumor sample. MSK-IMPACT testing on tumor identified a somatic frameshift mutation c.469delT (p.Tyr157 Thrfs*3) in *MLH1* exon 6 in the tumor. Loss of heterozygosity (LOH) of *MLH1* was observed in the patient's tumor sample (Figure 1b). MSK-IMPACT analysis of peripheral blood from the patient did not identify any clinically significant single-nucleotide variant (SNV) or indel variant in the *MLH1* coding region. In terms of copy number variation (CNV), we observed a −1.7 of fold change (p = 8.329e-11) for *MLH1* exon 7, suggesting a possible deletion of *MLH1* exon 7 (Table S1). However, the suspected deletion of *MLH1* exon 7 was not confirmed by MLPA (Figure S1). In a further analysis of the sequencing data, an insertion was revealed at c.588+8_588+9 in *MLH1* intron 7 by Integrative Genomics Viewer (IGV) (Figure 1c).

To investigate the nature and origin of the inserted fragment, PCR was performed with a combination of *MLH1* forward primer in exon 7 (F1) and reverse primer in exon 8 (R1) (Figure 1d). The control sample is expected to give a 460 bp band including exons 7–8, intron 7 and part of introns 6 and 8. The patient sample generates a 460 bp fragment from wild-type allele and a 570 bp band from the mutant allele (Figure 1d,e). Sanger sequencing confirmed the 460 bp band is the wild-type sequence while the 570 bp band includes an extra 115 bp

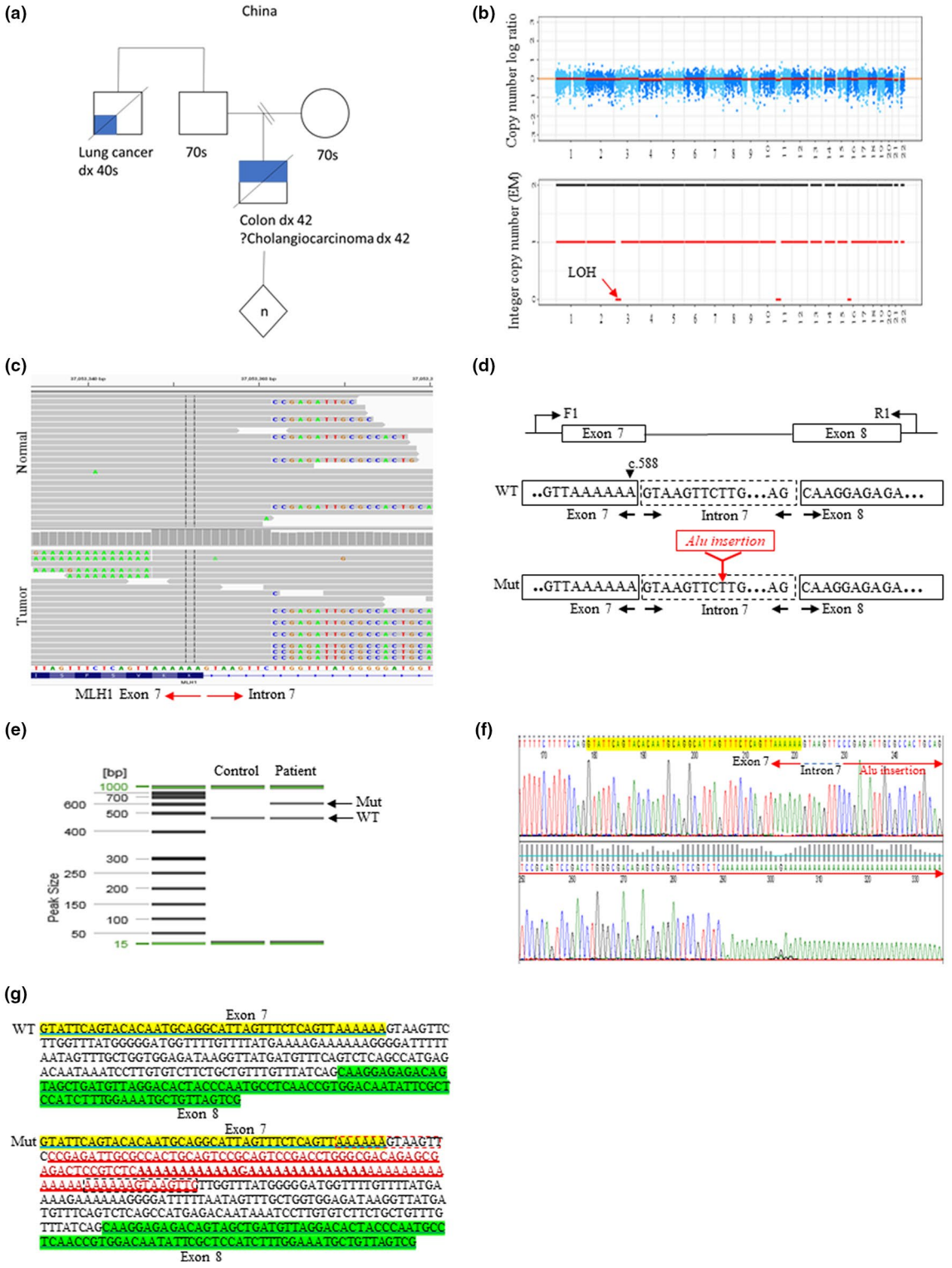

FIGURE 1 Identification of AluSx-like insertion in *MLH1* (NM_000249.3) intron 7. (a) Patient pedigree. The patient (indicated with the arrow) is a 42-year-old Chinese man was diagnosed with metastatic ascending colon. (b) FACETS show loss of heterozygosity (LOH) of *MLH1* in patient's tumor. (c) Insertion into c.588+8_588+9 of *MLH1* intron 7 by Integrative Genomics Viewer (IGV). (d) Schematic illustration of Alu insertion and PCR primer position. (e) A 460 bp fragment was generated from wild-type allele and a 570 bp band from mutant allele from the patient sample while the control only gives the 460 bp wild-type fragment. (f) Identification of the insertion by Sanger sequencing (Forward). (g) The sequence of the wild-type allele and the mutant allele of *MLH1* exon 7, intron 7, and exon 8

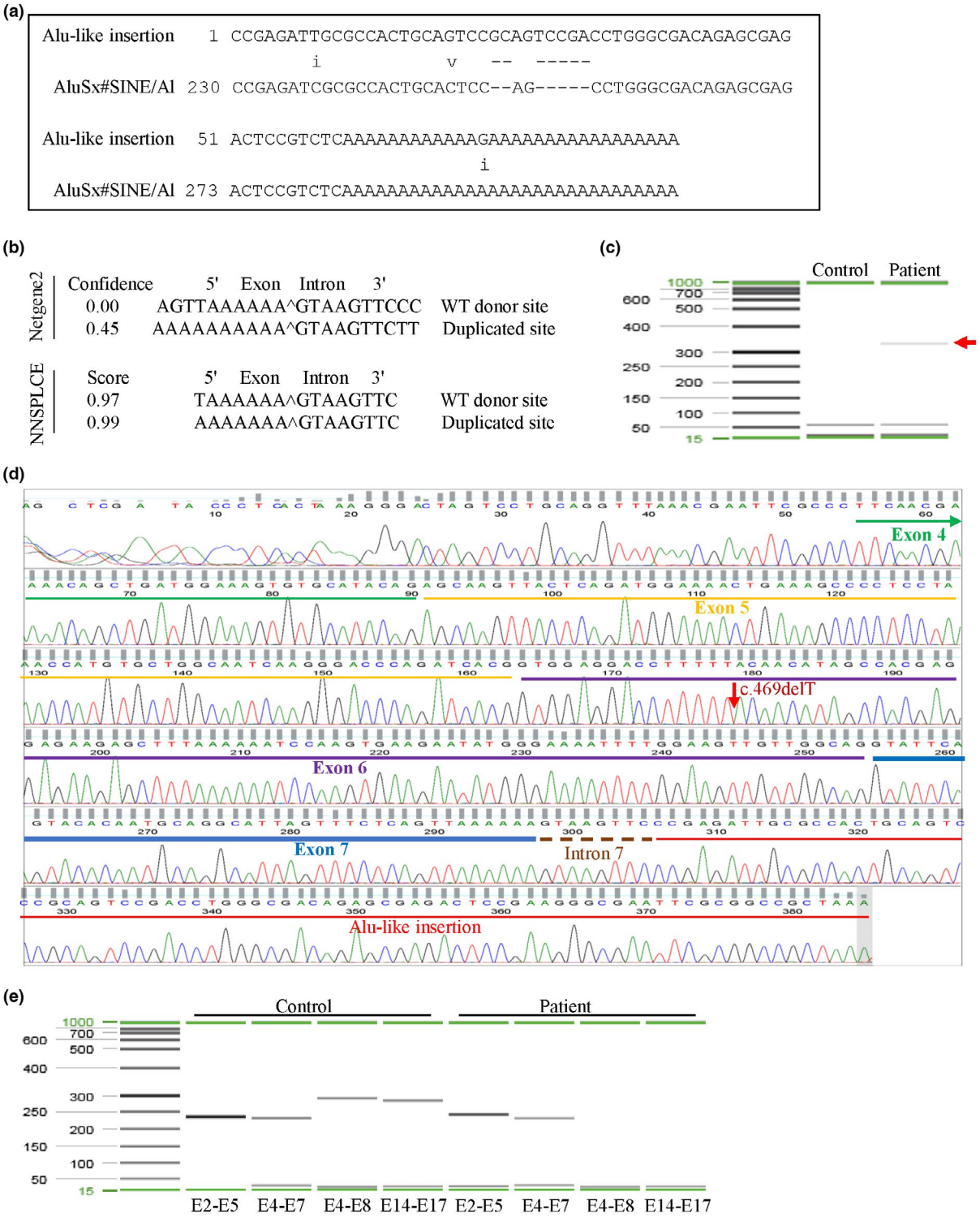

FIGURE 2 Disruption of transcription by the AluSx-like insertion in *MLH1* (NM_000249.3). (a) Sequence search shows that the insertion sequence is part of C terminal of AluSx#SINE/Al. (b) In silico tools Netgene2 and NNSPLCE predict higher possibility of usage of the duplicated donor site at the 3′ end of AluSx-like insertion than that of the WT donor site. (c) A 340 bp fragment was successfully amplified with the forward primer in *MLH1* exon 4 (Forward) and the reverse primer designed to 5′ end of the Poly (A) of the AluSx-like insertion by RT-PCR. (d) The AluSx-like insertion is introduced into *MLH1* mRNA. (e) Visible bands with the expected sizes were obtained with primer combinations of E2-E5, E4-E7, E4-E8, and E14-E17 in the control sample by RT-PCR. Primers E2-E5 and E4-E7 allowed us to visualize the expected RT-PCR bands in the patient sample. RT-PCR with primers E4-E8 and E14-E17 did not generate any product from the patient sample

insertion at the position of c.588+8_588+9 of *MLH1* intron 7 (Figure 1f,g).

In an analysis of the insertion sequence, the WT donor sequence from c.583 to c.588+8 (AAAAAAGTAAGTTC) of *MLH1* exon 7 is found to be duplicated at the 3′ end of the insertion, feature of a retrotranspositional event (Figure S2). The inserted sequence was queried in RepeatMasker Web Server and the result revealed that the inserted sequence is part of the C terminus of AluSx#SINE/Al (Figure 2a). Alu elements are typically approximately 300 bp in length. Although our insertion sequence is part of AluSx#SINE with only around 115 bp in length, it is featured with poly (A) and flanked with direct repeats duplicated from the sequences at the insertion site.

Insertions of Alu may disrupt splice sites or bind splicing regulators (Payer et al., 2019). Since the WT donor sequence from c.583 to c.588+8 (AAAAAAGTAAGTTC) of *MLH1* exon 7 is directly duplicated, which creates another donor site at the 3′ end of AluSx-like insertion, *in silico* analysis was performed to evaluate whether the AluSx-like insertion affects splicing. *In silico* tool Netgene2 predicts that the duplicated donor site at the 3′ end of AluSx-like insertion shows higher level of confidence while the WT donor site indicates loss of confidence (Figure 2b), suggesting that alternative usage of the duplicated donor site at the 3′ end of AluSx-like insertion, instead of the WT donor site at c.588 of *MLH1* exon 7. Another tool, NNSPLCE, also predicts higher score for the duplicated donor site than the WT donor site at c.588 (Figure 2b). Two other tools (SpliceSiteFinder and MaxEntScan) do not predict any difference (data not shown). *In silico* prediction is not very informative, at least for some alterations, and RNA analysis is needed to prove aberrant usage of the duplicated donor site.

Since LOH was observed in the tumor, we assumed the mutant allele would be detected by RT-PCR in the tumor tissue. To investigate whether the AluSx-like insertion affects splicing, RT-PCR was performed with RNA extracted from Formalin-Fixed Paraffin-Embedded (FFPE) tissue specimens. The forward primer of RT-PCR was in *MLH1* exon 4 (Forward) and the reverse primer was designed to bind to the 5′ end of the Poly (A) of the AluSx-like insertion. A 340 bp fragment was successfully amplified while the control sample did not generate a visible band (Figure 2c). To facilitate sequencing, this 340 bp fragment was cloned into a TA vector and subsequently sequenced with M13 tag from the TA vector. The sequencing result clearly demonstrated that the AluSx-like insertion is introduced into *MLH1* mRNA, confirming abnormal splicing with abolishment of the WT donor site of intron 7 (Figure 2d). MSK-IMPACT detected a somatic frameshift mutation c.469delT p.(Tyr157Thrfs*3) in *MLH1* exon 6 in the tumor. RT-PCR result confirmed the presence of the mutation c.469delT (Figure 2d). Both the Alu insertion (germline) and this c.469delT (somatic) are on the same allele.

To investigate whether the duplicated donor site (AAAAAAGTAAGTTC) at the 3′ end of AluSx-like insertion is alternatively used or the poly (A) of the AluSx-like insertion causes transcriptional termination, RT-PCR was performed with primer combinations from different *MLH1* exons. Forward primers bind to exons 2, 4, and 14 while reverse primers bind to exons 5, 7, 8, and 17. Visible bands with the expected sizes were obtained with primer combinations of E2-E5, E4-E7, E4-E8, and E14-E17 in the control sample by RT-PCR (Figure 2e). Primers E2-E5 and E4-E7 also allowed us to visualize the expected RT-PCR bands in the patient sample. However, RT-PCR with primers E4-E8 and E14-E17 did not generate any product from the patient sample, suggesting that transcription of the mutant allele was terminated in the Poly (A) region of the AluSx-like insertion (Figure 2e).

4 | DISCUSSIONS

In summary, we report here an AluSx-like element insertion into *MLH1* intron 7 which leads to abnormal transcriptional termination and causes Lynch syndrome. To our knowledge, this is the first report of Alu insertion into *MLH1* intronic regions outside the invariant region (+/- 1,2) of the splice consensus sequence. Our study broadens our understanding of Alu elements and their roles in cancer predisposition and allows other at-risk family members to undergo predictive germline testing to help identify family members who would benefit from increased cancer surveillance and/or risk-reducing measures.

This study demonstrated that when a well-designed genetic testing assay does not identify any pathogenic variants in individuals with strong personal and family history of Lynch Syndrome (or other cancer predisposition syndromes), the lab may need to pursue additional studies such as RNA sequencing of patient's normal and/or tumor (if available) as some noncoding variants may affect RNA splicing and gene function. The lab may also need to consider to optimize the bioinformatics pipeline to detect such challenging variants. We speculate that a substantial proportion of pathogenic variants might not be detected as most clinical assays target the coding regions of genes of interests and clinical labs restrict their analysis to coding regions and exon–intron boundaries. That could miss molecular events with clinical significance in deep intronic region, promoter region or 3′ UTR (Boyd et al., 2008; McConville et al., 1996; Mendes de Almeida et al., 2017; Nakamura et al., 2012; Smith et al., 2014; Vorechovsky, 2010).

ACKNOWLEDGMENT

We thank the administrative team (Silvana Ostafi, Marie Chamian, Stephanie Argueta, and Edyta B. Brzostowski) and Joshua Somar in the Diagnostic Molecular Genetics Laboratory at Memorial Sloan Kettering Cancer Center for their assistance on this project.

CONFLICT OF INTERESTS

Anna Varghese: Lilly (research); Verastem (research); BioMed Valley Discoveries (research); BMS (research); Silenseed (research); Dr. Stadler's immediate family member holds consulting/advisory roles in Ophthalmology with Allergan, Adverum Biotechnologies, Alimera Sciences, Biomarin, Fortress Biotech, Genentech/Roche, Novartis, Optos, Regeneron, Regenxbio, Spark Therapeutics. Dr. Zhang: Honoraria (Future Technology Research LLC, BGI, Illumina); Honoraria and Travel and accommodation expenses (Roche Diagnostics Asia Pacific). Family members hold leadership position and ownership interest of Shanghai Genome Center. The other authors do not have any conflict of interest.

AUTHOR CONTRIBUTIONS

LY designed and performed testing as well as drafted the manuscript. ESM, AV, MT, and ZKS provided saw the patient and returned testing results. LZ supervised the entire study. All authors reviewed the manuscript.

ORCID

Liying Zhang https://orcid.org/0000-0003-4517-0751

REFERENCES

Bonadona, V., Bonaiti, B., Olschwang, S., Grandjouan, S., Huiart, L., Longy, M., Guimbaud, R., Buecher, B., Bignon, Y. J., Caron, O., & Colas, C. (2011). Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome. *JAMA*, *305*, 2304–2310.

Boyd, C., Smith, M. J., Kluwe, L., Balogh, A., Maccollin, M., & Plotkin, S. R. (2008). Alterations in the SMARCB1 (INI1) tumor suppressor gene in familial schwannomatosis. *Clinical Genetics*, *74*, 358–366.

Cheng, D. T., Mitchell, T. N., Zehir, A., Shah, R. H., Benayed, R., Syed, A., Chandramohan, R., Liu, Z. Y., Won, H. H., Scott, S. N., Brannon, A. R., O'Reilly, C., Sadowska, J., Casanova, J., Yannes, A., Hechtman, J. F., Yao, J., Song, W., Ross, D. S., … Berger, M. F. (2015). Memorial sloan kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): A hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *The Journal of Molecular Diagnostics*, *17*, 251–264.

Cheng, D. T., Prasad, M., Chekaluk, Y., Benayed, R., Sadowska, J., Zehir, A., Syed, A., Wang, Y. E., Somar, J., Li, Y., Yelskaya, Z., Wong, D., Robson, M. E., Offit, K., Berger, M. F., Nafa, K., Ladanyi, M., & Zhang, L. (2017). Comprehensive detection of germline variants by MSK-IMPACT, a clinical diagnostic platform for solid tumor molecular oncology and concurrent cancer predisposition testing. *BMC Medical Genomics*, *10*, 33.

Cordaux, R., & Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nature Reviews Genetics*, *10*, 691–703.

Dewannieux, M., Esnault, C., & Heidmann, T. (2003). LINE-mediated retrotransposition of marked Alu sequences. *Nature Genetics*, *35*, 41–48.

Kim, J., Hu, C., Moufawad El Achkar, C., Black, L. E., Douville, J., Larson, A., Pendergast, M. K., Goldkind, S. F., Lee, E. A., Kuniholm, A., Soucy, A., Vaze, J., Belur, N. R., Fredriksen, K., Stojkovska, I., Tsytsykova, A., Armant, M., DiDonato, R. L., Choi, J., … Yu, T. W. (2019). Patient-customized oligonucleotide therapy for a rare genetic disease. *New England Journal of Medicine*, *381*(17), 1644–1652.

McConville, C. M., Stankovic, T., Byrd, P. J., McGuire, G. M., Yao, Q. Y., Lennox, G. G., & Taylor, M. R. (1996). Mutations associated with variant phenotypes in ataxia-telangiectasia. *American Journal of Human Genetics*, *59*, 320–330.

Mendes de Almeida, R., Tavares, J., Martins, S., Carvalho, T., Enguita, F. J., Brito, D., Carmo-Fonseca, M., & Lopes, L. R. (2017). Whole gene sequencing identifies deep-intronic variants with potential functional impact in patients with hypertrophic cardiomyopathy. *PLoS One*, *12*, e0182946.

Middha, S., Zhang, L., Nafa, K., Jayakumaran, G., Wong, D., Kim, H. R., Sadowska, J., Berger, M. F., Delair, D. F., Shia, J., Stadler, Z., Klimstra, D. S., Ladanyi, M., Zehir, A., & Hechtman, J. F. (2017). Reliable pan-cancer microsatellite instability assessment by using targeted next-generation sequencing data. *JCO Precision Oncology*, 1–17. https://doi.org/10.1200/PO.17.00084

Nakamura, K., Du, L., Tunuguntla, R., Fike, F., Cavalieri, S., Morio, T., Mizutani, S., Brusco, A., & Gatti, R. A. (2012). Functional characterization and targeted correction of ATM mutations identified in Japanese patients with ataxia-telangiectasia. *Human Mutation*, *33*, 198–208.

Payer, L. M., Steranka, J. P., Ardeljan, D., Walker, J., Fitzgerald, K. C., Calabresi, P. A., Cooper, T. A., & Burns, K. H. (2019). Alu insertion variants alter mRNA splicing. *Nucleic Acids Research*, *47*, 421–431.

Qian, Y., Mancini-DiNardo, D., Judkins, T., Cox, H. C., Brown, K., Elias, M., Singh, N., Daniels, C., Holladay, J., Coffee, B., Bowles, K. R., & Roa, B. B. (2017). Identification of pathogenic retrotransposon insertions in cancer predisposition genes. *Cancer Genetics*, *216–217*, 159–169.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). 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. *Genetics in Medicine*, *17*, 405–424.

Seth, S., Ager, A., Arends, M. J., & Frayling, I. M. (2018). Lynch syndrome - Cancer pathways, heterogeneity and immune escape. *The Journal of Pathology*, *246*, 129–133.

Smith, M. J., Wallace, A. J., Bowers, N. L., Eaton, H., & Evans, D. G. (2014). SMARCB1 mutations in schwannomatosis and genotype correlations with rhabdoid tumors. *Cancer Genetics*, *207*, 373–378.

Solassol, J., Larrieux, M., Leclerc, J., Ducros, V., Corsini, C., Chiesa, J., Pujol, P., & Rey, J. M. (2019). Alu element insertion in the MLH1 exon 6 coding sequence as a mutation predisposing to Lynch syndrome. *Human Mutation*, *40*, 716–720.

Teugels, E., De Brakeleer, S., Goelen, G., Lissens, W., Sermijn, E., & De Greve, J. (2005). De novo Alu element insertions targeted to a sequence common to the BRCA1 and BRCA2 genes. *Human Mutation*, *26*, 284.

Truninger, K., Menigatti, M., Luz, J., Russell, A., Haider, R., Gebbers, J.-O., Bannwart, F., Yurtsever, H., Neuweiler, J., Riehle, H.-M., Cattaruzza, M. S., Heinimann, K., Schär, P., Jiricny, J., & Marra, G. (2005). Immunohistochemical analysis reveals high frequency

of PMS2 defects in colorectal cancer. *Gastroenterology*, *128*, 1160–1171.

van der Klift, H. M., Tops, C. M., Hes, F. J., Devilee, P., & Wijnen, J. T. (2012). Insertion of an SVA element, a nonautonomous retrotransposon, in PMS2 intron 7 as a novel cause of Lynch syndrome. *Human Mutation*, *33*, 1051–1055.

Vorechovsky, I. (2010). Transposable elements in disease-associated cryptic exons. *Human Genetics*, *127*, 135–154.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

Fig S1

Fig S2

Table S1

How to cite this article: Li Y, Salo-Mullen E, Varghese A, Trottier M, Stadler ZK, Zhang L. Insertion of an Alu-like element in *MLH1* intron 7 as a novel cause of Lynch syndrome. *Mol Genet Genomic Med*. 2020;8:e1523. https://doi.org/10.1002/mgg3.1523