Insertion of an SVA Element in MSH2 as a Novel Cause of Lynch Syndrome

Ciyu Yang[1], Yirong Li[1], Magan Trottier[1], Michael Farrell[2], Vikas Rai[1], Erin Salo-Mullen[1], David Gallagher[2], Zsofia K. Stadler[1], Heleen van der Klift[3], and Liying Zhang[4]

[1]Memorial Sloan Kettering Cancer Center
[2]Mater Private Hospital
[3]Leiden University Medical Center
[4]UCLA

November 24, 2020

Abstract

Germline mutations in the DNA mismatch repair (MMR) genes cause Lynch syndrome (LS). Insertions of retrotransposons in MMR genes have been reported as a rare cause of LS. Here, we present a novel SINE-VNTR-Alu (SVA) insertion in exon 12 of MSH2 in an individual with early-onset colorectal cancer and strong LS family history. RT-PCR analysis indicated a larger aberrant MSH2 transcript in one of the family members. MSK-IMPACT next-generation sequencing testing and long-range PCR revealed an insertion in MSH2 exon 12 at the c.1972 position in an antisense orientation. The insertion was further characterized as an SVA element approximately 3 kb in length, belonging to the SVA_F1 family of retrotransposons.

Insertion of an SVA Element in *MSH2* as a Novel Cause of Lynch Syndrome

Ciyu Yang[1], Yirong Li[1], Magan Trottier[2], Michael P. Farrell[3], Vikas K. Rai[1], Erin Salo-Mullen[2], David J. Gallagher[3], Zsofia K. Stadler[2], Heleen M.van der Klift[4], Liying Zhang[1,5]

Departments of [1]Pathology,[2]Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065 USA

[3]Department of Cancer Genetics, Mater Private Hospital, Eccles St, Dublin 7, Ireland

Hosted file

`image1.emf` available at https://authorea.com/users/378238/articles/494798-insertion-of-an-sva-element-in-msh2-as-a-novel-cause-of-lynch-syndrome

[5]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095 USA

Corresponding author:

Liying Zhang, MD, PhD

Department of Pathology and Laboratory Medicine

David Geffen School of Medicine

University of California, Los Angeles

10833 Le Conte Ave, Los Angeles, CA 90095

Email: LiyingZhang@mednet.ucla.edu

ABSTRACT

Germline mutations in the DNA mismatch repair (MMR) genes cause Lynch syndrome (LS). Insertions of retrotransposons in MMR genes have been reported as a rare cause of LS. Here, we present a novel SINE-VNTR-Alu (SVA) insertion in exon 12 of *MSH2* in an individual with early-onset colorectal cancer and strong LS family history. RT-PCR analysis indicated a larger aberrant *MSH2* transcript in one of the family members. MSK-IMPACT next-generation sequencing testing and long-range PCR revealed an insertion in *MSH2* exon 12 at the c.1972 position in an antisense orientation. The insertion was further characterized as an SVA element approximately 3 kb in length, belonging to the SVA_F1 family of retrotransposons.

Keywords:

MSH2; Germline; SVA; Lynch Syndrome

Lynch syndrome (LS) is an autosomal dominant inherited disease and its prevalence is 1–3% in unselected colorectal or endometrial cancer patients (de la Chapelle, 2005). It is characterized by increased risks for early-onset tumor development, especially for colorectal cancer (CRC), endometrial cancer, ovarian cancer, and other extracolonic tumors such as hepatobiliary, urothelial, brain or central nervous system tumors, as well as sebaceous tumors (Cohen & Leininger, 2014). Lynch syndrome is caused by germline mutations in one of the mismatch repair (MMR) and *EPCAM* genes (Da Silva, Wernhoff, Dominguez-Barrera, & Dominguez-Valentin, 2016). Tumors from LS patients normally exhibit high microsatellite instability (MSI-H) and loss of expression of one or more MMR proteins (Boland, Koi, Chang, & Carethers, 2008). Substitutions, small insertion/deletions, large deletions/duplications, inversions (Liu et al., 2016; Mork et al., 2017; Rhees, Arnold, & Boland, 2014), as well as insertions of retrotransposon have been reported in the MMR genes as causes of LS (Peltomaki & Vasen, 2004; van der Klift, Tops, Hes, Devilee, & Wijnen, 2012).

Retrotransposons are DNA sequences that proliferate in the genome using an RNA intermediate and a 'copy-and-paste' retrotransposition mechanism. Retrotransposons can be subdivided into two groups distinguished by the presence or absence of long terminal repeats (LTRs). Retrotransposons without LTRs include Long Interspersed Elements 1 (LINE-1, L1), Alu elements (Short Interspersed Elements, SINE) and SVA (SINE-VNTR-Alu) elements (Cordaux & Batzer, 2009; Rebollo, Romanish, & Mager, 2012). Approximately 124 retrotransposon insertions associated with human disease have been previously reported (Hancks & Kazazian, 2016)

To date, ten (10) gross insertions larger than 20 base pairs (bp) have been recorded in MMR genes in the Human Genome Mutation database (HGMD Professional 2019.4). Five of these large insertions involved retrotransposons, four of which were Alu insertions with two in each of the *MLH1* (Leclerc et al., 2018; Solassol et al., 2019) and *MSH2* genes (Kloor et al., 2004; Marshall, Isidro, & Boavida, 1996) and one was an SVA insertion in *PMS2* (van der Klift et al., 2012). Up to now, no SVA insertion has been reported in *MSH2*. In this study, we report an insertion of an SVA element at c.1972 in exon 12 of *MSH2* as a novel cause of Lynch syndrome.

Our proband is a 49-year-old man who was diagnosed with colon cancer at age 43. A four-generation pedigree (Fig. 1) indicated that other family members were affected with early-onset colorectal cancer (CRC) under age 50. The proband's mother was diagnosed with metachronous endometrial and CRC and one maternal aunt was diagnosed with CRC at 50. One of the proband's brothers had colon polyps, and was subsequently diagnosed with a proximal colon cancer at age 54 which was MSH2 and MSH6 deficient on immunohistochemical (IHC) staining. Another brother was diagnosed with a screen detected colon cancer at age 38 which also demonstrated loss of expression of the MSH2 and MSH6 proteins by IHC. However, no mutation was identified through next generation sequencing (NGS) and the 10 Mb inversion in *MSH2* was not detected (Rhees et al., 2014).

Another maternal aunt of the index case was diagnosed with CRC at 35. Her daughter was diagnosed with endometrial cancer at age 38 which demonstrated MSI-H and loss of expression of MSH2 and MSH6 proteins by IHC. This family member was initially tested in 2007 for *MLH1* and *MSH2* sequencing and

large rearrangement in a reference laboratory and was identified to have an *MSH2* intron 12 rearrangement which was classified as variant of uncertain significance (VUS). Multiple family members affected with colon or endometrial cancer were tested and no mutation was identified, although tumor tissues of several individuals were tested and showed loss of MSH2 and MSH6 proteins with immunohistochemistry (IHC). A weak aberrant larger transcript was identified but not further characterized in lymphocyte RNA isolated from one of these family members who was affected with colorectal cancer (age 38) that showed loss of MSH2 and MSH6 expression (Fig. 2a). Additional Southern blot analysis on genomic DNA of the same patient indicated the presence of a 3 kb insertion, possibly a large LINE-1 or SVA insertion (Fig. 2b); restriction fragment analysis could narrow down the place of the insertion to a 1.45 kb region around MSH2 exon 12 (Fig. 2c). The same rearrangement was shown with Southern blot analysis in the genomic DNA from another more distantly related family member; this individual presented with endometrial cancer at age 34, and showed loss of MSH2 and MSH6 in tumor tissue. However, the type of retrotransposon and the exact genomic location of the insertion were not determined.

The proband was seen at the Clinical Genetics Service (CGS) at Memorial Sloan Kettering Cancer Center (MSKCC). Immunohistochemistry (IHC) analysis indicated loss of MSH2 and MSH6 proteins in the tumor. No *MSH2* inversion was detected. Given the strong family history of colon cancer, a colorectal multi-gene panel test (sequencing and large rearrangement analysis of *APC* , *EPCAM* (large rearrangement only), *MLH1, MSH2, MSH6, MUTYH, PMS2, POLD1, POLE* , with add-on genes: *PTEN, BRCA1, BRCA2*) was performed at the Diagnostic Molecular Genetics laboratory at MSK. Testing identified aberrant sequences before c.1954 and after c.1972 positions in exon 12 of *MSH2* (Fig. 3a). No other mutations or VUSs were identified in the remaining eleven genes analyzed. The copy number of *MSH2* exon 12 was normal based on our next-generation sequencing (NGS) analysis (Fig. 3b) which was confirmed by multiplex ligation-dependent probe amplification (MLPA) (Fig. 3c), indicating that the aberrant sequence was probably not due to a genomic deletion or duplication of the coding region of *MSH2* .

To investigate the nature and origin of the abnormal sequence, long-range (LR) PCR, was performed on genomic DNA from the patient using the TaKaRa LA PCR Kit according to the manufacturer's protocol, an M13-tagged forward primer located in intron 11 (5'- GTA AAA CGA CGG CCA GT GGGTTTTGAATTC-CCAAATG - 3') and an M13-tagged reverse primer in intron 12 (5'- CAG GAA ACA GCT ATG AC AAAACGTTACCCCCACAAAG-3'). One band about 400 bp in length was present in negative controls (Fig. 4a). Another band of a larger size (~3 - 4 kb) was observed in the patient but absent in the negative controls (Fig. 4a). The larger aberrant fragment was extracted and sequenced with M13 forward and reverse primers. Sequence analysis of the extracted aberrant fragment revealed a targeted duplication of 19 bps, with the location of the insertion at c.1972, and part of the inserted sequence (660 bp) in an antisense orientation with respect to the *MSH2* transcription direction (Fig. 4b, 4c). The Repeat Masker (http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker) indicated the inserted sequence belongs to an SVA element. To map the inserted sequence, we performed a BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi, human genome GRCh38.p12 reference, Annotation Release 109). A total of 27 hits covered almost all chromosomes, except chromosome 16, 18 and Y. The first alignment showed close to 100% homology between the 559 inserted sequence bp (except one nucleotide) and a region on chromosome 3 (Chr3: 48,210,600 - 48,211,258). The Repeat Masker indicated that an SVA repeat was present at this location.

The sequence from 48,210,600 to 48,220,000 in chromosome 3 was retrieved from the NCBI database and sequencing primers were designed based on the retrieved sequence. Apart from a short gap in the VNTR region that could not be sequenced due to the repetitive structure, we were able to decipher a total of 2,937 bp inserted sequence without including the polyA tail in the antisense strand (Fig. 5a). The inserted sequence identified in our proband starts with a guanine nucleotide followed by a 295 bp exon 1 of the *MAST2* gene, an Alu-like element (37 bp), an approximately 2.2 kb VNTR region with tandem repeats ranging from 37 to 54 bp long, a SINE-R region (492 bp), the putative polyadenylation signal AATAAA, and a long polyA tail. The inserted sequence is followed by a target site duplication (TSD) of 19 bps (*MSH2* c.1954_1972) at the 3'end of the insertion (Fig. 4c, 5a). Interestingly, the inserted sequence is very close to the sequence

3

from 48,210,600 to 48,213,111 in chromosome 3, except that the insertion has a longer VNTR (Fig. 5a, 5b). Further analysis characterized it as a human specific SVA subfamily of retrotransposons termed SVA_F1 that contains a *MAST2* 5' transduction group and is a fusion of *MAST2* exon 1 containing CpG island and a 5'-truncated SVA (Bantysh & Buzdin, 2009; Damert et al., 2009; Hancks, Ewing, Chen, Tokunaga, & Kazazian, 2009) (Fig. 5c, 5d). Seventy-six members have been identified in the SVA_F1 subfamily in the human genome. In 96% of SVA_F1 members, the SVA element insert starts with a guanine residue (Bantysh & Buzdin, 2009) and the SVA_F1 insertion in this case also starts with a 'G' (Fig. 4c, 5a).

The SVA insertion in *MSH2* exon 12 likely occurred through LINE-1-mediated retrotransposition as it exhibits several classical features of this process (Hancks et al., 2009; Raiz et al., 2012) as shown in Fig. 4c and 5a: (1) insertion at consensus LINE-1 endonuclease cleavage site 5'-TTTT/AA-3' (where "/" denotes the cleavage site); (2) the presence of a direct repeat TSD of 19 bp in length, within the size range of 4–20 bp that is typical for LINE-1 mediated retrotranspositions; (3) a long polyA tail preceded by the putative polyadenylation signal AATAAA; and (4) presence of 5' transducing and truncation, a structural variation encompassing more than 8% of all SVA elements in the human genome (Damert et al., 2009; Raiz et al., 2012; Wang et al., 2005).

In summary, we describe here for the first time an SVA insertion into the coding sequence of *MSH2* mediated by LINE-1 protein machinery. Precise location of SVA insertion and determination of the specific SVA sequence in the *MSH2* gene are important for cancer management to guide genetic testing of family members and potentially preimplantation genetic testing. Furthermore, cancer affected family members identified to have Lynch Syndrome may further benefit from immune checkpoint inhibitors which are FDA-approved for MMR deficient and MSI-H tumors, the hallmark of Lynch Syndrome associated tumors. Therefore, identification and characterization of the SVA elements and their roles in cancer predisposition genes paved the path for genomic precision medicine and cancer prevention and therapy.

Data Sharing and Data Availability

We have shared our data with the scientific community and deposited this variant into the ClinVar database. Here is the link: https://submit.ncbi.nlm.nih.gov/subs/clinvar_wizard/SUB8577511/overview

ID: SUB8577511

Compliance with Ethical Standards:

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

Conflict of Interest: Z.K. Stadler reports that an immediate family member serves as a consultant in the field of ophthalmology for Adverum Biotechnologies, Genentech/Roche, Novartis, Neurogene, Gyroscope Tx, Optos Plc, Regeneron, RegenexBio, and 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 interests of Scipher Health.

Figure legends

Fig. 1 Patient pedigree. The patient described here is a 49-year-old man who was diagnosed with colon cancer at 43. Nine family members were affected with colorectal cancer and six members were diagnosed with endometrial cancer in this family.

Fig. 2 RT-PCR and Southern blot analysis. a) Agarose gel electrophoresis of the RT-PCR of *MSH2* exon 6-16 on mRNA of short-term cultured lymphocytes treated with (+chx) or without (-chx) cycloheximide. A very faint band, approximately 200 bp larger than the normal RT-PCR product, is visible in patient's RNA sample treated with CHX. The PCR is not quantitative; the faintness of the larger band is possibly caused by problematic amplification of a GC-rich aberrant transcript, or diminished expression of the mutant allele. b) Southern blot gels of genomic DNA of the patient (two independent DNA samples PAT_D1 and PAT_D2)

4

and two negative controls (CTR1 and CTR2) digested with four different restriction enzymes and hybridized with an MSH2 exon 10-16 cDNA probe (M2-3'b). All enzymes show an approximately 3 kb larger extra fragment in the patient's DNA in addition to the normal wildtype fragment. c) Restriction fragment analysis and approximate localization of the insertion in the *MSH2* gene.

Fig. 3 MSK-IMPACT and MLPA results. a) IGV shows aberrant sequence in exon 12 of *MSH2* . b) CNV plots of the patient showed no deletion or duplication in MSH2 exon 12. c) MLPA results confirmed no CNV change in *MSH2* of the patient.

Fig. 4 Long-range PCR and sequencing results after gel extraction. a) LR-PCR products run on 1% agarose gel. An extra band were observed in the patient, but not in controls. b) Electropherogram showing the site of insertion. c) Partial inserted sequences in red in anti-sense orientation with respect to the *MSH2* gene transcription direction. Target site duplication is highlighted in yellow. DNA sequence in the first row matches with those in reverse direction in b.

Fig. 5 SVA insertion in exon 12 of the *MSH2* gene in the anti-sense direction. a) Final inserted sequence identified. Sequences that are not present on Chromosome 3 are highlighted in gray. " ..." indicates sequences that were not deciphered. b) Sequence on chromosome 3 (GRCh38.p12 Primary Assembly from 48,210,600 to 48,213,111.). The sequences that are not present in the SVA insertion in MSH2 are highlighted in yellow. The sequences present in the SVA insertion in MSH2 but not on Chromosome 3 are left in blank. c) Schematic view of full length SVA_F1 (not in scale). d) Schematic representation of the SVA element inserted in *MSH2* exon 12 (not in scale).

REFERENCES

Bantysh, O. B., & Buzdin, A. A. (2009). Novel family of human transposable elements formed due to fusion of the first exon of gene MAST2 with retrotransposon SVA.*Biochemistry (Mosc), 74* (12), 1393-1399. doi:10.1134/s0006297909120153

Boland, C. R., Koi, M., Chang, D. K., & Carethers, J. M. (2008). The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch syndrome: from bench to bedside. *Fam Cancer, 7* (1), 41-52. doi:10.1007/s10689-007-9145-9

Cohen, S. A., & Leininger, A. (2014). The genetic basis of Lynch syndrome and its implications for clinical practice and risk management. *Appl Clin Genet, 7* , 147-158. doi:10.2147/TACG.S51483

Cordaux, R., & Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nat Rev Genet, 10* (10), 691-703. doi:10.1038/nrg2640

Da Silva, F., Wernhoff, P., Dominguez-Barrera, C., & Dominguez-Valentin, M. (2016). Update on Hereditary Colorectal Cancer. *Anticancer Res, 36* (9), 4399-4405. doi:10.21873/anticanres.10983

Damert, A., Raiz, J., Horn, A. V., Lower, J., Wang, H., Xing, J., . . . Schumann, G. G. (2009). 5'-Transducing SVA retrotransposon groups spread efficiently throughout the human genome. *Genome Res, 19* (11), 1992-2008. doi:10.1101/gr.093435.109

de la Chapelle, A. (2005). The incidence of Lynch syndrome. *Fam Cancer, 4* (3), 233-237. doi:10.1007/s10689-004-5811-3

Hancks, D. C., Ewing, A. D., Chen, J. E., Tokunaga, K., & Kazazian, H. H., Jr. (2009). Exon-trapping mediated by the human retrotransposon SVA. *Genome Res, 19* (11), 1983-1991. doi:10.1101/gr.093153.109

Hancks, D. C., & Kazazian, H. H., Jr. (2016). Roles for retrotransposon insertions in human disease. *Mob DNA, 7* , 9. doi:10.1186/s13100-016-0065-9

Kloor, M., Sutter, C., Wentzensen, N., Cremer, F. W., Buckowitz, A., Keller, M., . . . Gebert, J. (2004). A large MSH2 Alu insertion mutation causes HNPCC in a German kindred.*Hum Genet, 115* (5), 432-438. doi:10.1007/s00439-004-1176-9

Leclerc, J., Flament, C., Lovecchio, T., Delattre, L., Ait Yahya, E., Baert-Desurmont, S., . . . Buisine, M. P. (2018). Diversity of genetic events associated with MLH1 promoter methylation in Lynch syndrome families with heritable constitutional epimutation. *Genet Med, 20* (12), 1589-1599. doi:10.1038/gim.2018.47

Liu, Q., Hesson, L. B., Nunez, A. C., Packham, D., Williams, R., Ward, R. L., & Sloane, M. A. (2016). A cryptic paracentric inversion of MSH2 exons 2-6 causes Lynch syndrome.*Carcinogenesis, 37* (1), 10-17. doi:10.1093/carcin/bgv154

Marshall, B., Isidro, G., & Boavida, M. G. (1996). Insertion of a short Alu sequence into the hMSH2 gene following a double cross over next to sequences with chi homology.*Gene, 174* (1), 175-179. doi:10.1016/0378-1119(96)00515-x

Mork, M. E., Rodriguez, A., Taggart, M. W., Rodriguez-Bigas, M. A., Lynch, P. M., Bannon, S. A., . . . Vilar, E. (2017). Identification of MSH2 inversion of exons 1-7 in clinical evaluation of families with suspected Lynch syndrome. *Fam Cancer, 16* (3), 357-361. doi:10.1007/s10689-016-9960-y

Peltomaki, P., & Vasen, H. (2004). Mutations associated with HNPCC predisposition – Update of ICG-HNPCC/INSiGHT mutation database. *Dis Markers, 20* (4-5), 269-276. doi:10.1155/2004/305058

Raiz, J., Damert, A., Chira, S., Held, U., Klawitter, S., Hamdorf, M., . . . Schumann, G. G. (2012). The non-autonomous retrotransposon SVA is trans-mobilized by the human LINE-1 protein machinery. *Nucleic Acids Res, 40* (4), 1666-1683. doi:10.1093/nar/gkr863

Rebollo, R., Romanish, M. T., & Mager, D. L. (2012). Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annu Rev Genet, 46* , 21-42. doi:10.1146/annurev-genet-110711-155621

Rhees, J., Arnold, M., & Boland, C. R. (2014). Inversion of exons 1-7 of the MSH2 gene is a frequent cause of unexplained Lynch syndrome in one local population. *Fam Cancer, 13* (2), 219-225. doi:10.1007/s10689-013-9688-x

Solassol, J., Larrieux, M., Leclerc, J., Ducros, V., Corsini, C., Chiesa, J., . . . Rey, J. M. (2019). Alu element insertion in the MLH1 exon 6 coding sequence as a mutation predisposing to Lynch syndrome. *Hum Mutat, 40* (6), 716-720. doi:10.1002/humu.23725

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. *Hum Mutat, 33* (7), 1051-1055. doi:10.1002/humu.22092

Wang, H., Xing, J., Grover, D., Hedges, D. J., Han, K., Walker, J. A., & Batzer, M. A. (2005). SVA elements: a hominid-specific retroposon family. *J Mol Biol, 354* (4), 994-1007. doi:10.1016/j.jmb.2005.09.085
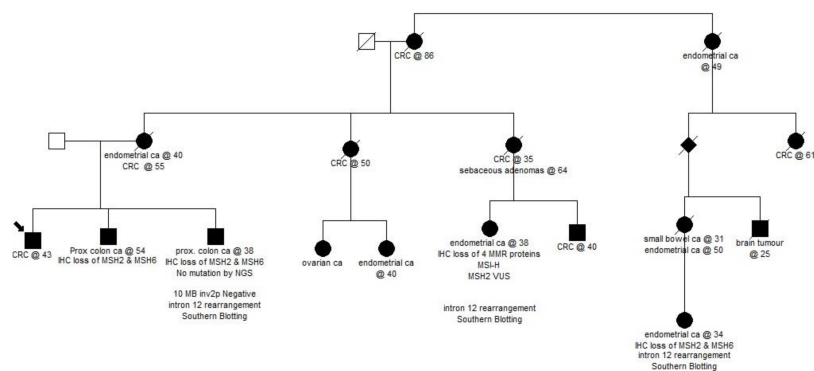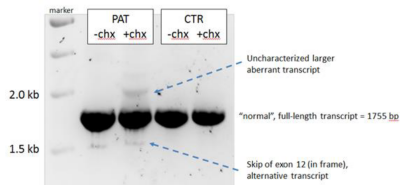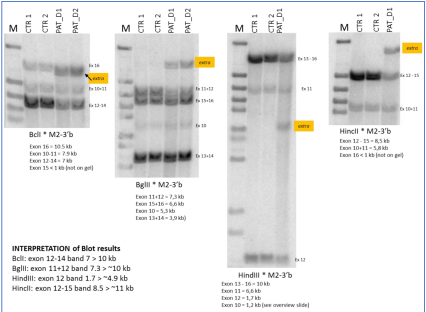
Figure 1

6

Figure 2a

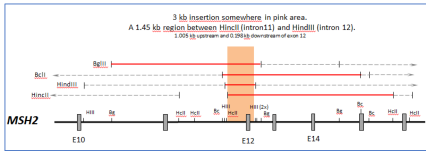PAT -chx +chx   CTR -chx +chx

marker

2.0 kb
1.5 kb

Uncharacterized larger aberrant transcript
"normal", full-length transcript = 1755 bp
Skip of exon 12 (in frame), alternative transcript

Figure 2b

BclI * M2-3'b
Exon 16 - 10.5 kb
Exon 10-11 - 7.9 kb
Exon 12-14 - 7 kb
Exon 15 < 1 kb (not on gel)

BglII * M2-3'b
Exon 11+12 = 7.3 kb
Exon 15+16 = 6.6 kb
Exon 10 - 5.5 kb
Exon 13+14 - 3.9 kb)

HincII * M2-3'b
Exon 12 - 15 = 8.5 kb
Exon 10-11 = 5.8 kb
Exon 16 < 1 kb (not on gel)

INTERPRETATION of Blot results
BclI: exon 12-14 band 7 > 10 kb
BglII: exon 11+12 band 7.3 > ~10 kb
HindIII: exon 12 band 1.7 > ~4.9 kb
HincII: exon 12-15 band 8.5 > ~11 kb

HindIII * M2-3'b
Exon 13 - 16 = 10 kb
Exon 11 - 6.6 kb
Exon 12 - 1.7 kb
Exon 10 - 1.2 kb (see overview slide)

Figure 2c

3 kb insertion somewhere in pink area.
A 1.45 kb region between HincII (intron11) and HindIII (intron 12).
1.005 kb upstream and 0.296 kb downstream of exon 12

MSH2
E10   E12   E14

Figure 3a

c.1954   c.1972

Figure 3b

Probe level SCNA calls

Show 10 entries   Search: msh2

Chr	Pos	Cytoband	Gene	Exon	Fold Change	p_value
2	47630328	2p21	MSH2	exon1	-1.02594775767975	0.999154554066191
2	47693794	2p21	MSH2	exon10	-1.12389718194711	0.999154554066191
2	47698101	2p21	MSH2	exon11	-1.09059401557981	0.999154554066191
2	47702161	2p21	MSH2	exon12	-1.13866846795812	0.925367687476622
2	47703503	2p21	MSH2	exon13	1.02367985312265	0.999154554066191
2	47705408	2p21	MSH2	exon14	-1.06595649510598	0.999154554066191
2	47707832	2p21	MSH2	exon15	-1.04591962527742	0.999154554066191
2	47709915	2p21	MSH2	exon16	-1.04106018000927	0.999154554066191
2	47635537	2p21	MSH2	exon2	-1.02683984145373	0.999154554066191
2	47637230	2p21	MSH2	exon3	-1.01490450884825	0.999154554066191

Figure 3c

Blue1

31	MSH2-E01b	146.9	1.041
32	MS2-E01b	189.6	0.973
33	MS2-E01c	216.9	0.939
34	MS2-E02	160.4	1.026
35	MS2-E03	172.8	1.003
36	MS2-E04	183.9	0.956
37	MS2-E05	211.4	1.074
38	MS2-E06	231.8	0.990
39	MS2-E07	250.2	1.043
40	MS2-E08	271.6	0.956
41	MS2-E09	292.9	1.023
42	MS2-E10	310.1	0.917
43	MS2-E11	331.2	0.965
44	MS2-E12	346.9	1.020
45	MS2-E13		
46	MS2-E14	390.1	0.888
47	MS2-E15	407.0	1.001
48	MS2-E16	425.9	1.046
49	MS2-07-1	265.0	
50	MS2-07-2	317.0	-1

Figure 4a

Negative 1   Negative 2   The patient   Blank

bp   bp
5,000   5,000
4,000   3,000
3,000   2,000
2,000   1,600
1,000   1,000
850   850
650   650
500   500
400   400

Figure 4b

Forward direction

c.1954   c.1972

Reverse direction

Exon 12   c.1972   c.1954   Inserted sequence

Figure 4c

Intron 12   Exon 12   c.1972   c.1954   Duplicated sequence

cctaagaaccaggttttttacCAGTAATGATGTGGAACATCTGTTTATCTTTTCAAAGTATACGTCATTAGG

Inserted sequence
GAGAAGGGTCGAAGGCGCCGCGGGCTGGGGTCGGTGGCTTAGGGAGCCCGTCTGGCCATGGTGGCCGCGGCTG
GTGGTTGGCGCGGCGGCGCTGACGCCCGCGGGCCCCAGCTGCAGATATGAAGCGGAGCCGCTGCGCGGACCGAC
CGCAGCCGCCGCCGCCGACCGCCAGGAGGGATGGAGTTCAGCGGGCAGCGGAGCTGTCTCAGTCTTTGCCGCCG
CGCCGGCGAGCGCCGCCCGGGAGGCAGCGGCTGGAGGAGCGGACGGGCCCCGCGGGGCCCGAGGGCAAGGAG
CAGCCGCCTGCCTTGGCCTCCCAAAGTGCCGAGATTGCAGCCTCTGCCCGGCTGCCACCCCGTCTGGGAAGTGAG
GAGTGTCTCTGCCTGGCCGCCCATCGTCTGGGATGTGAGGAGCCCCTCTGCCTGGCTGCCCAGTCTGGAAAGTGA
GGAGCGTCTCCGCCCGGCCGGCCGCCATCCCATCTAGGAAGTGAGGAGCGCCTCTTCCCAGCCGCCATCACATCTAGGA
AGTGAGGAGCGTCTCTGCCCGGGCCGCCCATCGTCTGAGATGTGGGGAGCGCCTCTGCCCCGCCGCCCCATCTGGG
ATGTGAGGAGCGCCTCTGCCCGGCCGAGACCCCGTCTGGGAGGTGAGGAGCGTCTCTGCCCGGGCCGC............
.................AAAAAAAAAAAAA(A)n

CAAAGTATACGTCATTAGG

c.1972   c.1954

Exon 12
AATAAATGCAATTTCATCTTGAACTTCAACACAAGCATGCCTGGATGCTTTTAATATAATTCTTCCTTGTCCTTTCTCC
AAAATGGCTGGTCGTACATATGGAACAGGTGCTCCATTTGACACGTGAGCAAAGCTGACAACAGCATCTAGCTGA
GCTAACACATCATTGAGTGTCTGCATTGGTTCTACATAGCctgtataaaataaaaacagaaaatgtacacaggaatactgaataa
taaacagatataagccgtaaaaatacataatcccccccagtgatcgc
Intron 11

Figure 5a **Figure 5b**

Figure. 5c

Figure. 5d